Vol. 37, No. 2 March, 1930

PSYCHOLOGICAL REVIEW PUBLICATIONS

Psychological Review

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY
MADISON BENTLEY, CORNELL UNIVERSITY (*J. of Exper. Psychol.*)
S. W. FERNBERGER, UNIV. OF PENNSYLVANIA (*Bulletin*)
W. S. HUNTER, CLARK UNIVERSITY (*Index*)
RAYMOND DODGE, YALE UNIVERSITY (*Monographs*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY, *Business Editor*

ADVISORY EDITORS

R. P. ANGIER, YALE UNIVERSITY; MARY W. CALKINS, WELLESLEY COLLEGE; JOSEPH JASTROW, UNIVERSITY OF WISCONSIN; C. H. JUDD, UNIVERSITY OF CHICAGO; ADOLF MEYER, JOHNS HOPKINS UNIVERSITY; W. B. PILLSBURY, UNIV. OF MICHIGAN; C. E. SEASHORE, UNIVERSITY OF IOWA; G. M. STRATTON, UNIV. OF CALIFORNIA; MARGARET F. WASHBURN, VASSAR COLLEGE; JOHN B. WATSON, NEW YORK; R. S. WOODWORTH, COLUMBIA UNIVERSITY.

CONTENTS

The Future of Psychophysics: F. M. URBAN, 93.

The Use of Equality Judgments in Psychophysical Procedures: SAMUEL W. FERNBERGER, 107.

Some Properties of Fechner's 'Intensity of Sensation': H. M. JOHNSON, 113.

The Present Status of Experimental Studies on the Relationship of Feeling to Memory: H. MELTZER, 124.

The Reliability Coefficient as a Per Cent, with Application to Correlation between Abilities: ROBERT CHOATE TRYON, 140.

Intelligence Tests of Immigrant Groups: CARL C. BRIGHAM, 158.

Science and Cult: W. S. TAYLOR, 166.

Motive as a Concept in Natural Science: FLOYD H. ALLPORT, 169.

Discussion:

Experimental Extinction and Negative Adaptation: A. L. WINSOR, 174.

Measures of Musical Talent: A Reply to Dr. C. P. Heinlein: CARL E. SEASHORE, 178.

A Reply to Professor Bentley: J. F. DASHIELL, 183.

PUBLISHED BI-MONTHLY
FOR THE AMERICAN PSYCHOLOGICAL ASSOCIATION
BY THE PSYCHOLOGICAL REVIEW COMPANY
PRINCE AND LEMON STS., LANCASTER, PA.
AND PRINCETON, N. J.

Entered as second-class matter July 13, 1897, at the post-office at Lancaster, Pa., under Act of Congress of March 3, 1879

Psychological Review Publications

OF THE

American Psychological Association

EDITED BY

HOWARD C. WARREN, PRINCETON UNIVERSITY (*Review*)
RAYMOND DODGE, YALE UNIVERSITY (*Monographs*)
SAMUEL W. FERNBERGER, UNIV. OF PENNSYLVANIA (*Bulletin*)
W. S. HUNTER, CLARK UNIVERSITY (*Index*)
MADISON BENTLEY, CORNELL UNIVERSITY (*J. of Exp. Psych.*)
HERBERT S. LANGFELD, PRINCETON UNIVERSITY, *Business Editor*

WITH THE COOPERATION OF
MANY DISTINGUISHED PSYCHOLOGISTS

PSYCHOLOGICAL REVIEW

containing original contributions only, appears bimonthly, January, March, May, July, September, and November, the six numbers comprising a volume of about 500 pages.

PSYCHOLOGICAL BULLETIN

containing critical reviews of books and articles, psychological news and notes, university notices, and announcements, appears monthly, the annual volume comprising about 720 pages. Special issues of the BULLETIN consist of general reviews of recent work in some department of psychology.

JOURNAL OF EXPERIMENTAL PSYCHOLOGY

containing original contributions of an experimental character, appears bimonthly, February, April, June, August, October, and December, the six numbers comprising a volume of about 500 pages.

PSYCHOLOGICAL INDEX

Is a compendious bibliography of books, monographs, and articles upon psychological and cognate topics that have appeared during the year. The INDEX is issued annually in June, and may be subscribed for in connection with the periodicals above, or purchased separately.

PSYCHOLOGICAL MONOGRAPHS

consists of longer researches or treatises or collections of laboratory studies which it is important to publish promptly and as units. The price of single numbers varies according to their size. The MONOGRAPHS appear at irregular intervals and are gathered into volumes of about 500 pages.

ANNUAL SUBSCRIPTION RATES

Review: $3.50 (Foreign, $3.75)
Journal: $6.00 (Foreign, $6.25)
Bulletin: $6.00 (Foreign, $6.25)
Monographs: $6.00 per volume (Foreign, $6.30)
Psychological Index: $1.00

Current Numbers: Review or Journal, $1.00; Bulletin, $.60

Review and Bulletin, $10.00 (Foreign, $10.50)
Journal and Bulletin, $11.00 (Foreign, $11.50)
Review and Journal, $10.00 (Foreign, $10.50)
Review, Bulletin, and Journal, $15.00 (Foreign, $15.75)
Review, Bulletin, Journal, and Index, $18.00 (Foreign, $18.75)

Subscriptions, orders, and business communications may be sent direct to

PSYCHOLOGICAL REVIEW COMPANY,
PRINCETON, N. J.

VOL. 37, No. 2 March, 1930

THE PSYCHOLOGICAL REVIEW

THE FUTURE OF PSYCHOPHYSICS[1]

BY F. M. URBAN

Brünn, Czechoslovakia

I do not claim to be a prophet or to possess revelations. I will try to read the future from the things of today. A single work hardly can give you a clue as to what is going to happen, but the tendencies in the work of present-day investigators are likely to persist and may give a hint as to future development. A tendency is an idea which has got hold of a man. These ideas are not clearly expressed, and but dimly perceived even by those who follow them. Recognize them, and you have the directive power of research.

Mark that there is no way of foretelling the result of any particular investigation. An experiment is an intelligent question put to nature and it is no sign of intelligence to inquire for what one knows already. Nature often is a reluctant witness and you need all your resourcefulness to press the important questions. In the end she usually surrenders to persistent effort and it rarely happens that experimental research undertaken by a competent investigator remains without reward. It is easy to predict brilliant results for all tendencies which lead to experimentation.

It is equally easy to foretell the inevitable clash of conflicting opinions. Divergent views on general questions result in different interpretations of facts and the representatives of opposite tendencies are thrown in contact with one another

[1] Read at the Ninth International Congress of Psychology held at New Haven, Conn., September, 1929.

on innumerable points. Nothing short of a miracle can prevent disputes arising out of these disagreements, and we all know that this miracle does not happen.

If nature is enigmatic, investigators are human and you may gauge their motives to some extent. Nature's lines are mysterious, but grand. The motives of the investigator are intelligible enough, but their variety puzzles us. An unsolved problem or a suspected inconsistency of results are the most frequent motives for starting an investigation, that panacea against intellectual discomfort. You all know the predicament of the investigator who has grappled with a problem unsuccessfully. He has tried every way of approach he can think of, and finds himself checked everywhere. He gives up in despair and leaves the problem unsolved, hoping to return to it later, when an improvement in method or an insight gained elsewhere may have opened a new line of attack. Sometimes one has to wait too long for this happy turn. Other problems, more promising than the first, gain your interest and your first love gradually wanes and disappears from your mind. You would not return to it, even if success were tolerably certain.

This happens to the individual and this happens in science as well. Many a problem looms up large before the eyes of one generation and—though unsolved—hardly finds any interest in the next. It is this human element which makes the progress of science so puzzling and so erratic at times. Sooner or later somebody will take up the old problem, but you do not know whether a tendency will find its representatives in the next generation. You, therefore, must not be too optimistic about the date when your predictions will come true.

Let us see whether we can find in my own work any well pronounced tendencies and let us judge their prospects to persist. I am responsible for papers and books on a variety of topics, but they all center about psychophysics. Sometimes the connection seems loose, but I never wrote on a problem which did not first confront me in psychophysics. Neither did anyone of the side issues supersede my interest

in it, not even the calculus of probability to which, under the stress of circumstances, I had to devote most of my work in recent years. It is strange that there should be so many ramifications of the work in this apparently limited field, and I never would have suspected it in the beginning of my career. I feel very much like John Wilkes who in 1799 said in his book on pen-cutting: "Had I not taken to the profession of pen-cutting, it never would have occurred to me that there should be so much to say on making pens."

The papers usually start with some experimental facts, followed by mathematical considerations. What is behind this curious mixture of experience and mathematics, and what part do the formulæ play? Let me express my answer in three short rules: Keep close to experience; add as little of your own as possible; if you have to add something, be mindful to give an account of every step you take.

I do not anticipate any objections against the first rule, to let the facts speak for themselves. Some people might even try to outdo me by insisting that science must contain facts and nothing else. Unfortunately this cannot be done. You cannot deal with statistical data without some sort of mathematical treatment and the use of a formula may have unforeseen consequences. It may imply suppositions which are entirely justified in other fields of work, but absolutely unsuitable to the peculiar conditions of psychology.

The immediate result of psychophysical experiments are numbers of relative frequency of the different judgments. They are the facts, but they do not mean much, if they remain isolated. Indeed, you cannot even conclude that a repetition of the experiments will give the same or at least similar results. You first have to get an idea of the conditions which brought about these results, and of their prospect to remain constant. Applying the results to experiments not yet made is an inductive generalization in the sense of Keynes, and must be supported in some way. If you want to draw conclusions from numbers, you are bound to use mathematics, whether you like it or not.

You cannot do statistical work without formulæ, and if

you could, you would not do it, because mathematics is too powerful an instrument. Only you have to be careful in its use, for otherwise you might find yourself in trouble. It rarely pays to put implicit faith in a method without knowing the restrictions of its usefulness. It holds of mathematics too, that you must not add too much of it to your results. Let me illustrate this by an example.

Lagrange's formula is the simplest method of representing the psychometric functions. It merely implies that the probabilities of the different judgments are continuous functions of the stimulus, and this supposition will hardly be challenged. The formula adds nothing to the data and it gives merely the mathematical representation of the experimental findings. It has, however, two shortcomings. 1. The calculated results for the stimuli in the lower and upper part of the scale do not agree with our ideas about the course of the psychometric functions. Some results are impossible, since they give for the probabilities values above unity or below zero, and others show fluctuations while the psychometric functions should increase or decrease uniformly. 2. The formula does not take account of the errors of observation. It looks like a waste of energy to build up a complicated expression of which one knows beforehand that it cannot be final, because it is affected by the contingencies of an individual set of experiments. Empirical data are never exact and their inconsistency necessitates further mathematical treatment. This is the reason why you cannot leave your data such as they are but have to add something of your own.

The problem of eliminating the errors is solved in an admirable way by the use of Tchebytcheff's series or by the series of Φ-functions. They take account of the errors, but add nothing to the experimental data. The latter series is called by some authors Charlier's series, while others refer it to Bruns. We will not go into the historical question as to whom the credit is due—it may be that the priority belongs to Hermite and to neither of these two men—and call the series by the indifferent name of the functions which are used in it. Let me call here your attention to a peculi-

arity of scientific usage. An author is held responsible for any errors or mistakes of which he may be guilty, but his lasting achievements remain unnamed. The fault is personal, but merit anonymous. It makes the acquisition of truth look like the work of an infallible brotherhood, into which the individual contributions are merged, and this seems to suit well the impersonal character of science.

Both series are perfectly general and may be adapted to any set of results. They do not imply definite hypotheses about the psychometric functions and in this they resemble Lagrange's formula, but they are superior to it by taking account of the limited accuracy of the observations. These series probably give the most satisfactory solution. Considerable progress was made recently by Lufkin through the use of the series of Φ-functions, but of this I will speak later.

A definite hypothesis about the psychometric functions is implied, if you adjust your results according to some particular function. The $\Phi(\gamma)$-hypothesis is the best known example. It seems that Pearson's formulæ also belong to this class. In spite of Thomson's brilliant work the use of the Pearsonian formulæ in psychophysics is not as well understood as it might be. It is not proven that the expressions obtained by Pearson's method are the best obtainable, so that there might be other expressions which give smaller sums of the squares of the residuals. It is not certain that Pearson's results are the best obtainable. It may be that such proof will be forthcoming, but until now it is not at hand and this is a slight blemish on Pearson's beautiful work.

Psychophysical work with Pearson's formulæ must remain difficult. These expressions are rather complicated and in psychophysics we have to deal with their integrals. To this come two practical difficulties. The first is, that the first and last cells of the pseudohistogram slant. The second refers to the fact that the number of stimuli which can be used conveniently is rather limited. Seven or nine stimuli are probably the limit for practical work, otherwise the experiments become unmanageable. I am not fully con-

vinced that these formulæ may be profitably applied to pseudohistograms of such limited dimensions.

It seems to me that these formulæ are not as general as the two series aforementioned. It is not proven that all distributions arising from the binomial expansion may be represented by one of Pearson's formulæ. I am informed that Professor Mirimánoff and his pupils are working on this question and it may be that we will have definite information on this question within short time. Should the answer be in the negative, as I surmise, Pearson's formulæ are but slightly superior to the $\Phi(\gamma)$-hypothesis.

There are a variety of questions clustering about this one point of the application of the Pearsonian formulæ to psychophysics. They are all straightforward problems admitting of an answer in one way or the other. Their solution requires nothing but straight thinking and the application of methods which are known to every competent worker.

I am fully aware of all the difficulties to which the adjustment of experimental results by definite hypotheses about the psychometric functions is open, and I have done my share in explaining them to the public. I fully appreciate the force of Boring's argument on the futility of the search after a generalized psychometric function. Boring certainly is right in this, that there exists no expression which applies to all cases. It may be, however, that one particular expression gives satisfactory results in a certain class of experiments. It would be rather remarkable, if this should not be so. The psychometric functions may be different for trained and untrained subjects, or for optic, acoustic, and lifted weight experiments, but it is incredible that there should exist absolute irregularity. The search for such regularities is legitimate and I do not believe that it can be given up, for it is one of the most fascinating problems. Boring's argument is valuable in so far as it bids us to be on our guard and not to dogmatize. I trust to live long enough to see a definite advance in this direction, but I will ever be mindful of Boring's warning and regard these functions, no matter what their success may have been, merely as convenient

expressions which do good work in their own province. Outside of it they may be less suitable, and it will not surprise me if they fail entirely in representing some new material.

This example may have explained what is meant by the caution, to add to experience as little as possible. Before following up this rule into a wider field, I want to go on record with a prediction. The time is not distant when professional mathematicians will take up psychophysical problems. In some respects this is a cheerful outlook, for their specialistic training will enable them to view our problems from a new standpoint and to apply methods of reasoning with which we are less familiar. In this respect nothing but good can come from such a collaboration, but I must say a word of warning to our prospective collaborators. Mathematical knowledge alone does not guarantee success in psychophysical work. You need in addition, or rather to start with, an intimate knowledge of the problems and practical acquaintance with the work. Had I to choose as collaborator between a fair experimentalist and a perfect mathematician, I would take the experimentalist every time. He will do practical work and not turn my psychophysical problems into mathematical exercises. It is not proven either that a pure mathematician necessarily has a deeper insight into our problems. Bruns and Borel, two men of outstanding ability, tried their hands at psychophysics and achieved but indifferent success. Neither of them found the correct formula for the weight of the observation equations in the method of constant stimuli, which is all the more surprising since this is a purely mathematical affair.

You all know that there are certain differences of opinion between my friend Dr. Wirth and myself. Their source is that rule of mine, that you should add to experience as little as possible. My most important difference with Dr. Wirth refers to the notion of the threshold. Wirth starts from this notion and builds up on it his entire system of formulæ. I regard the threshold as a superfluous hypothesis, in itself not more than a metaphorical expression for the fact that a stimulus is perceived if it is above the threshold, and not

perceived if it is below the threshold. This hypothesis is superfluous, as is seen by the fact that all the formulæ may be obtained without it. I start from the relative frequencies and resolve psychophysics into a study of the psychometric functions and their parameters. The proof of my assertion that the notion of the threshold is not needed for the foundation of psychophysics is complete, since Wirth's formulæ are exclusively based on the frequencies of the judgments. To this Wirth replies that his formulæ are meaningless to him, if they are not interpreted in terms of the threshold.

Our discussion is unique in some respects. We both stated our cases and then kept silent for almost twenty years. Wirth was the last to talk. His answer did not convince me, but I felt that further argumentation against such a personal view was useless. To my mind a statement does not become less significant by cutting out metaphors and putting in a straightforward way. A metaphor forces you to constant circumlocutions and the presentation must needs become awkward. I am convinced that Wirth's presentation may be given in a much shorter space. This is not due to any cleverness of mine, for none of us can hope to surpass Dr. Wirth in ability, but to the simple fact that a straightforward presentation is shorter than one burdened with metaphors.

I do not imagine that I can convert Dr. Wirth to my view, neither is it likely that I will change my opinion. It is a deep-going conflict in which the parties concerned cannot agree, but it is a great satisfaction to me that this difference of opinion never reacted on our personal relations and that we collaborated on more than one occasion. Our case is before the jury and the judgment on the merits of our views on the threshold hypothesis is reserved to our fellow-workers, present and future. None of the parties can hope for a draw and the matter is too vital to them to take the possibility of defeat into serious consideration. I therefore take it for granted that the decision will be in my favor and give you an outline of the future textbook on psychophysics, which does not start from the threshold-hypothesis.

It will begin with a very careful analysis of introspection

and its importance in psychophysical experimentation. Much has been done in this respect already and there are men present to whom we owe some notable contributions, but the results might still be improved upon. I merely call attention to the fact that we know very little about the introspective standard, by which a subject decides on the judgment to be given on the comparison of two stimuli. It would be easy to enumerate several other points on which introspective evidence is deplorably defective, but it will suffice to mention the lack of information on this crucial point, in order to show how much introspective work remains to be done.

It was pointed out that there is a certain behavioristic tendency in my psychophysical views. This is true to this extent that, according to my mind, the results must be treated as if nothing were known about them but the observed numbers of relative frequency. Statistical data, to be significant, must refer to well-defined conditions. One of the decisive conditions in psychophysical work—probably the most important of all—is the mental state of the subject. This is directly given by introspection, but inaccessible in any other way. It is absurd to keep the physical conditions scrupulously constant and to study the reactions of the subject without bothering about his mental state. The experimenter must first define the introspective conditions, under which the experiments are performed. The subject must have sufficient introspective training to judge during the experiments, whether his mental state was such as required by the instruction. He must reject experiments in which he did not comply with the instructions. Letting an experiment pass unchallenged, must be equivalent to the statement that introspectively nothing can be said against it. In a word: You have to collect your statistical data under conditions which are physically as well as introspectively well defined and constant. Introspection is needed for planning and performing the experiments, but when these are over all further work consists in a purely statistical treatment of the observed numbers of relative frequency.

The next chapter of our imaginary textbook will bring a searching scrutiny of the numbers of relative frequency. The question to be answered is this: Under which circumstances and how far may these numbers be regarded as empirical determinations of the underlying unknown probabilities. It is certain that the most careful control of all physical and mental conditions will be required in order to materialize the case of constant probabilities, and every new insight into the influences to which the judgments are subjected, will bring a new refinement of procedure. Let me illustrate this by an example.

Fernberger observed that the order in which the pairs of stimuli are presented is not immaterial to the result. The judgments are influenced by the experiments immediately preceding. Leaving aside the question as to the mechanism of this influence, we may roughly say that it is of the type of inhibition and reinforcement. It is strongest immediately after the stimulation and decreases rather rapidly as the time elapsed since the first stimulation increases. The very obvious conclusion is that one has to insert an interval between consecutive experiments. Had Fernberger's observation been made twenty years ago, I would have inserted an interval of rest—*e.g.*, one of four metronome beats—between each pair of liftings and thereby probably improved my results. Experimental technique is nothing fixed. What is perfect today, may be open to criticism in a year and unsatisfactory in ten years. From this follows the necessity to repeat the fundamental experiments from time to time, as the progress of experimental technique may suggest improvements.

The next chapter of the future textbook on psychophysics will contain the theoretical presentation of the methods. There will be no fundamental change at this point and the matter will be practically the same as today. Of course there will be many improvements in detail and the presentation will be at once more elegant and more cogent, but the essential results will remain unchanged. I believe that the foundations of psychophysics are well laid and that we have

to anticipate enlargements and additions, but no fundamental changes.

At the same time the number of problems accessible to exact treatment will be much larger than it is today. It is an easy matter to point one's finger at gaps in our present knowledge. I do not speak of questions, the solution of which is entirely beyond the ken of present day methods, but of problems which are known to exist and of which it is tolerably certain that they may be solved by our methods. I mention the method of mean error, the theory of which is in a very unsatisfactory state. This method is most interesting from the theoretical point of view, for it uses the Gaussian coefficient of precision as a measure of sensitivity. This, apparently, is a departure from the methods of constant stimuli and just perceptible differences, which use the interval of uncertainty. Should there be no relation between the coefficient of precision and the interval of uncertainty, psychophysics could not be a homogeneous system.

The theory of errors of observations must be based on psychophysics, and we know already that there exists a simple relation between the Gaussian coefficient and the interval of uncertainty as determined by the methods of constant stimuli and by the method of just perceptible differences. This point is not of theoretical interest only, but of paramount importance to the credit of psychophysics among the other sciences, for the answer will settle the puzzle which the method of least squares has offered for more than a hundred years.

We want to know as much as possible about the method of mean error. It should be analyzed in the same way as the method of just perceptible differences, *i.e.* an experimental material should be produced which may be worked out by this method as well as by the methods of constant stimuli and of just perceptible differences. The difficulty is not a mean one, to make a pun on the name of our problem, but it is after all merely one of experimental technique, and there is no reason why some clever person should not devise a procedure by which this may be done.

The chapter on methods will also include the analysis of many processes which until now escaped serious attention of psychologists. I mention the processes by which the oculist fixes eye-glasses or by which the ear-specialist determines the acuity of hearing. Both are very reliable and admit of a very fair degree of accuracy, as any person with astigmatism can testify. They take much less time than a determination of sensitivity by the method of constant stimuli, even if Fernberger's most approved improvements or Spearman's shortcuts are used. Processes of such evident merits should not be neglected. I am the last to attach an exaggerated value to practical utility, but it seems to me that theory and practice ought to be studied alongside. They both will profit by it.

Let me say a word on the old argument of the futility of too great refinement in methods. Some people insist that our results are rather inexact and admit of the very simplest kind of mathematical treatment only. They tolerate the use of footrules only. It is sensible to insist that we should have simple formulæ for treating results of small extent. We have such formulæ, and Wirth has given some notable examples of them. It is not sensible, however, to treat results of short and long series of experiments in the same way. We have sets of psychophysical data which in reliability and accuracy compare very favorably indeed with those of other statistical sciences, and the energy spent in collecting the data is wasted if we do not get as much out of them as there is in them. The conscience of the psychophysicists is clear in this respect, for it is part of their work to test the reliability of their results and not to go in their calculations beyond the limit warranted by the data.

Short formulæ as well as refined formulæ have their use in the proper places. A textbook on psychophysical methods may be compared to a shop in which you can buy all kinds of knives. Razors as well as pocket-knives are good things, but if you try to use them promiscuously you will be disappointed. The attempt at shaving with your pocket-knife will have disastrous, though foreseeable, results, and if you

sharpen a pencil with a razor, you will break it and the pencil will remain as dull as ever.

The formal description of the methods will be followed by the presentation of the results obtained by them. Here my power of prediction stops and I will not venture further. My paper might finish here, but I should like to make an observation about different ways of achieving progress, which was suggested to me by two outstanding contributions of recent years. I mean the work of Culler and of Lufkin. These men achieve progress in two entirely different ways.

Culler's work begins with experimentation which must be called model. I expressed this opinion before, but public discussion of minor differences of opinion apparently obscured in the eye of the reader my high esteem for Culler's work. Public discussion of disputable points easily gives the impression that the merit of the whole work is put in question. This most decidedly is not the case of Culler's work. I still believe that I am right on the points criticized and I am ready to fight on this issue, but this does not prevent me from saying that his work has set a new standard in experimentation and that it will have a far-reaching influence. Culler had the idea how experiments in a new field ought to be conducted, and his success is the due reward of his faithful work.

Lufkin's line of attack is different. He had the idea of using a method of calculation, spoken of in a haphazard way many times before, but not ever really worked out. His success was inevitable from the moment he decided to start his investigation. His discovery lay ready for anyone who had the right idea and actually worked it out. There were several men who might have made this discovery, but Lufkin alone actually made it. I speak feelingly on this topic, for I am one of the men in question. Was it laziness or a kind of mental inertia which let this opportunity slip away from me?

In the beginning of scientific advance stands the idea. It may refer to new experimentation, to new methods of research, or to a new way of looking at things. People are

bound to have different ideas and out of their clash, commonly called discussion, truth and new insight must arise. A Greek philosopher has said: πόλεμος πατὴς παντῶν. Let me translate it into the vernacular thus: Discussion is the mother of progress.

[MS. received September 17, 1929]

THE USE OF EQUALITY JUDGMENTS IN PSYCHO-PHYSICAL PROCEDURES *

BY SAMUEL W. FERNBERGER

University of Pennsylvania

The problem of the categories of judgment to be used in psychophysical experiments is almost as old as the methods themselves. So far as I know, every worker in this field has been willing to admit the categories of 'greater' and of 'less' judgments. Those of 'guess' and of 'don't know' have been variously handled and are beside the scope of the present paper. It is the question of the 'equality' judgment which I wish to raise.

In very early psychophysics, the equality judgment had its place alongside of the other categories. It was admitted without question—this is the common sense point of view. Indeed, in the early form of the Method of Just Perceptible Differences, the notion of equality was the starting point for two of the four experimental procedures for determining the positive and negative perceptible differences. And here is where the trouble started. It was not possible to find any stimulus between the threshold values which would remain stable and which would give a judgment of equality every time one came back to it. This can be readily understood inasmuch as these values of stimulus must fall within what Urban has so appropriately called the Interval of Uncertainty where the judgment is so unstable. G. E. Müller solved the difficulty by devising his modification of this Method of Just Perceptible Differences in a way which eliminated the notion of equality as a starting point for any of the experimental procedures but in a way which still retains the equality judgment.

The handling of the equality judgment has led to various

* Read at the Ninth International Congress of Psychology held at New Haven, Conn., September, 1929.

procedures in the past. Fechner [1] suggested that the equal judgments be divided in half and that one half be added to the right and one half to the wrong cases. This procedure was followed by Fullerton and Cattell.[2] Merkel,[3] followed by Kraepelin,[4] suggested that the equal judgments be not only excluded from the calculations but also from the records. Jastrow,[5] Higier [6] and Sanford [7] contended that there should be no equal judgments and that when the subject does not know whether the judgment should be greater or less, he should guess. Warner Brown [8] in 1910 even asserts that, if the subject is forced to give a judgment of greater or less, he can do so. It is doubtful whether the numerical results of any of these procedures would vary greatly from Fechner's frank admitting of the equality judgments to be then divided for purposes of calculation.

Titchener [9] pointed out the original reasons why the equal judgments were first excluded. He shows that it was because three classes of judgment presented difficulties in mathematical treatment that were unsurmountable at that time. But these difficulties have since been solved and a further exclusion of judgments of this class for that reason, he calls unscientific.

Urban,[10] Wirth [11] and Thompson [12] have supplied the mathematical procedures for handling the psychometric functions for the classes of greater, less, and equal. Indeed, if

[1] G. Fechner, Elemente der Psychophysik, 1889, I, 72, 94.

[2] G. S. Fullerton & J. McK. Cattell, The perception of small differences, 1892, 59.

[3] J. Merkel, Das Psychophysische Grundgesetz in Bezug auf Schallstärken, *Philos. Stud.*, 1887, 4, 131.

[4] E. Kraepelin, Zur Kenntniss der psychophysischen Methoden, *Philos. Stud.*, 1890, 6, 496.

[5] J. Jastrow, A critique of psycho-physic methods, *Amer. J. Psychol.*, 1888, 1, 282.

[6] H. Higier, Experimentelle Prüfung der psychophysischen Methoden im Bereiche des Raumsinnes der Netzhaut, *Philos. Stud.*, 1891, 7, 247.

[7] E. S. Sanford, A course in experimental psychology, 1898, 357.

[8] W. Brown, The judgment of difference, 1910.

[9] E. B. Titchener, Experimental psychology, 1905, II, Part 2, 290.

[10] F. M. Urban, Die psychophysischen Massmethoden als Grundlagen empirischer Messungen, *Arch. f. d. ges. Psychol.*, 1909, 15, 261–355; 1910, 16, 168–227.

[11] W. Wirth, Psychophysik, 1912.

[12] G. H. Thompson, The accuracy of the phi-gamma process, *Brit. J. Psychol.*, 1914, 7, 44.

one defines the limens in terms of a probability of 0.5, the equality judgments become of primary importance. For Urban has shown that the area bounded by the psychometric function for the equality judgments and the abscissa is equal to the rectangle with the interval of uncertainty as base and with unity as height. In this figure the height, unity, is fixed by definition. Hence any increase in the number of equality judgments, with a corresponding increase in the size of both of these areas, must result in a wider separation of the threshold values and a corresponding increase in the size of the interval of uncertainty. Of course, the limen is a mere statistical threshold and it is chosen arbitrarily —the 0.5 limen is the easiest to calculate in the first place and besides seems more logical by definition than any other. But, if any other threshold value were chosen, the relation between the area under the curve of the psychometric function for the equality judgments and the separation of the thresholds remains. This whole question is raised again because in certain recent literature there again seems to be the tendency to eliminate the equality judgments.

What is the trouble with them—why are they so difficult to handle experimentally? As early as 1907, Frank Angell [13] pointed out that some subjects give more equality judgments than others and that, to quote him, "this difference corresponds to the differences between deliberate and impulsive temperaments." A number of years ago, I had an observer who, in the course of an extended experiment, failed to give a single equality judgment.[14] Thus his upper and lower thresholds coincided and his interval of uncertainty was zero. This implies that this observer could judge the difference between a standard and a comparison stimulus with a probability of 0.5 or more even though the difference between the two stimuli were infinitely small.

In the course of this experiment, it became evident that the usual instructions to judge in terms of three categories,

[13] F. Angell, On judgment of 'like,' *Amer. J. Psychol.*, 1907, 18, 253–260.

[14] S. W. Fernberger, On the elimination of the two extreme intensities of the comparison stimuli in the method of constant stimuli, PSYCHOL. REV., 1914, 21, 335–355.

motivate the average subject to look for a difference. The average observer searches for a difference and, failing to find one, he then and only then, turns to equality, which in many cases is a sort of an admission of weakness. In a later introspective study,[15] if one may still mention introspection in this so-called behavioristic age, it became evident that the equality judgments almost invariably involved doubt, that they required verification on the part of the observer and that they are given with a longer reaction time than are the judgments of difference. And this from highly trained observers. Indeed, the attributive patterns for the difference judgments and for the equality judgments were so different that they were treated in separate sections in this paper. These findings were not only verified by S. S. George [16] but he also found that only judgments of greater, equal, and less could be given under a constant attitude. Reiter [17] found that it was as valid to give judgments of equal or not equal as it was to obtain those of greater or not greater, and of less or not less. The curves of the psychometric functions were similar in all three cases to those obtained under the ordinary three-category instruction.

I thus tried out the thing experimentally.[18] To one group of observers the usual instructions were given which tend to motivate the subject to look for a difference. To another group of observers were given instructions which emphasized the equality judgment—which assured the observer that the giving of an equality judgment was not a sign of weakness. And for the second group, we obtained an average interval of uncertainty which was over 60 per cent greater than that for the first group of observers.

This is the real difference between the equality and the

[15] S. W. Fernberger, An introspective analysis of the process of comparing, Psychol. Monog., 1919, 26, No. 6, pp. 161.

[16] S. S. George, Attitude in relation to the psychophysical judgment, *Amer. J. Psychol.*, 1917, 28, 1–37.

[17] F. H. Reiter, The influence of instruction on the formation of judgments in lifted weight experiments, Philadelphia, 1917. Thesis of the University of Pennsylvania, pp. 21.

[18] S. W. Fernberger, The effect of the attitude of the subject upon the measure of sensitivity, *Amer. J. Psychol.*, 1914, 25, 538–543.

difference judgments. The difference judgments are much more 'objective' in nature. The equality judgments are more dependent on the attitude of the observer, on the form of instruction, on the observer's understanding of instructions and, as Angell pointed out, to an extent, at least, on his basic temperament. Here, as elsewhere, it is difficult to handle situations which are subjective in nature, especially without the frequent introspective report of the observer. Such a report would limit the use of the psychophysical measurement methods to the normal adult observer who is further trained in systematic introspection. And the psychophysical methods have been handled in a truly behavioristic fashion—in a way to which even Watson could not take exception. The stimuli have been presented to the observers who have reacted with the verbal report of a single word.

Difficult as they are to handle experimentally, can we afford to eliminate the equality judgments? As early as 1906, Titchener [19] said: "If we are dealing with mind as mind presents itself to us for examination, we cannot ignore these judgments." George's and my own introspective studies show that judgments of true subjective equality occur frequently for all observers.

What do we wish to obtain by the psychophysical methods after all? The old idea that we are measuring the sensitivity of a particular sense organ has been abandoned. We now recognize that we are measuring the sensitivity of the entire psychophysical organism—his sense organs, his concentration, his attitude, his acceptance and understanding of instructions, his degree of practice, and what not besides—to a given stimulus situation. It is recognized that all of these things may change and it is for this reason that we obtain a large number of observations which are then treated statistically, rather than a single observation.

It is evident, then, that we need a program for further experimental research to determine the influences effecting such judgments. This would lead to a standardized set of conditions for psychophysical experimentation. We now

[19] E. B. Titchener, Experimental psychology, 1905, II, Part 2, 290.

have well standardized conditions covering the more objective phases of most psychophysical work. What we need is similar standardization of the more difficult and more illusive subjective phases.

[MS. received September 26, 1929]

SOME PROPERTIES OF FECHNER'S 'INTENSITY OF SENSATION'

BY H. M. JOHNSON

Mellon Institute of Industrial Research, University of Pittsburgh

In my recent critique (*4*) of the logic of Fechner's system of psychophysics (*2*, *3*), I presented certain considerations which indicate that comparison of the intensities of any two Fechnerian sensations [1] is meaningless. To save space, I omitted mention of some implications of Fechner's defining equation (*Massformel*) which seemed to be rather obvious, and which yield the same conclusion as the propositions that I did give. The omission however has been criticized, on pedagogical grounds which may be valid. I shall therefore exhibit the implications here, in the light of some elementary considerations by which they are to be interpreted. The latter will be mentioned first.

By the logical criterion expounded by Bridgman (*1*), every concept is to be defined in terms of a set of operations.[2] In general, if two distinct sets of operations are considered, they define two distinct concepts, which, to avoid ambiguity, should be called by different names. To illustrate the method of defining by operations rather than by properties, we may recall two distinct concepts which are indifferently styled *length*. The first concept is defined by applying standard

[1] The term *Fechnerian sensation* is used herein to denote a sensation, whether directly observable or not, whose intensity is expressed by Fechner's *Massformel*.

[2] Bridgman gives the best formal exhibition that I have seen of a method actually followed by Mach (*5*), by Poincaré, and by many others, of defining their scientific concepts in terms of operations instead of properties. He also emphasizes the disasters that may attend a departure from the operational method. I consider his book an outstanding contribution to the economy of universal science, no less than to that of physics. By the operational criterion, many physical concepts prove, of course, to be undefined, and all propositions concerning them prove to be meaningless. The case is far worse in the biological, educational and social sciences; it is particularly so in psychology. A rigorous application of the method to the analysis of psychological 'problems' would show a huge proportion to be spurious, and would forestall much unnecessary and irrelevant experimentation.

measuring rods, such as meter sticks, to bodies whose lengths are required, and by counting the necessary applications. The result may be called *tactual length*. The second concept is defined by a more complicated set of operations, employed in situations like the following. The distance c between two points A and B is required. The distance b from point A to C, a point not in the line AB, is given, in terms of tactual length. The three points A, B and C thus determine a triangle, of which the side b, opposite the angle B, is known. By means of such an instrument as a surveyor's transit, one measures the angle of intersection at point A of the light-beams BA and CA; and the angle of intersection at point C of the light-beams AC and BC. If the instrument can be located at point B the angle of intersection of light-beams AB and CB is also measured, and the three sets of readings 'adjusted' so that the sum of the three angles shall equal 180°. Otherwise one simply defines angle B as 180° less the sum of angles A and C. This is done on the *assumption* that the light-beams follow the 'straight lines' of Euclidean space. The test of the assumption is impossible, except through a range that is very small as compared to astronomical distances, and even within this range, the result of the test is attended by a finite uncertainty. The distance c between points A and B is now calculated from a formula of Euclidean geometry which relates the length of an unknown side to that of a given side and to the adjacent angles, thus:

$$c = \frac{b \cdot \text{Sin } C}{\text{Sin } [180° - (A + C)]}.$$

This operation is quite different in kind from the one which distinguishes *tactual length*, and its result is to be distinguished by the name *optical length*. The two concepts are not identical. Within macroscopic but terrestrial distances, and subject to a finite uncertainty, the two sets of results may be treated as numerically equivalent, and interchangeable within the range for which the equivalence has been demonstrated by simultaneous comparison. Such a comparison cannot be applied to astronomical distances. It is meaningless to ask

whether the equivalence holds for astronomical space, until the tactual method can be applied thereto. To assume that it does is to extrapolate, far beyond the range of experience, a relationship which has been observed to hold through only a restricted range, and subject within that range to an appreciable uncertainty. As Bridgman emphasizes, one of the most important lessons to be drawn from the present revolution in physics is that concepts should be defined in terms of actual operations, so that one may remain prepared to encounter novelties whenever the range of experience is about to be increased.

We may now examine the concept of Fechnerian intensity of sensation by the operational criterion. This concept, which is designated by the symbol γ, is exhibited in the defining equation known as Fechner's *Massformel*, thus:

$$\gamma \equiv k \cdot \text{Log}\ (\beta/b),$$

in which β represents the magnitude of the physical stimulus to which the sensation is referred, while b is that value of β which under the conditions of the experiment produces a 'threshold-effect' upon the responses of the organism; k being a constant of proportionality depending on one's choice of a system of logarithms and on the unit in terms of which γ is to be expressed. Both β and b are expressions of kinetic Energy or of its equivalent in Work. Both are defined by *physical* operations. The operation which results in the selection of b from the various values of β consists, in Fechner's Method of Right and Wrong Cases, for example, in *classifying* the responses as *right* or else as *wrong*, and then in *counting* the 'right' responses which correspond to each value of β that was employed. That value of β is called b which is found by actual count, or by interpolations in the actual count, to yield (say) 84% of responses classified as 'right.' Having obtained β by *physical* measurement, and having selected b from a series of physical measurements by a *social* criterion; and having then performed the *mathematical* operations of division, taking a logarithm, and multiplying, one achieves the concept γ. Its meaning is the operations that

create it. It is often discussed, however, as if it were interchangeable with another concept, which we may call the *intensity of introspectional sensation*, and denote by the symbol I. The question is, whether $I = \gamma$. To give an answer that satisfies the operational criterion we must first specify the operations by which I is determined. If they prove to be identical with the operations which determine γ, then I and γ are not mere numerical equivalences, they are identities; and the meaning of I is simply $k \cdot \text{Log}\,(\beta/b)$. But if I means some other set of operations, we must perform them, and compare the result with the simultaneous value of γ. If the two sets of results, taken in simultaneous pairs, agree within the required standard of precision, we may say that within the range of the experiment, and subject to a specified degree of uncertainty, they are *mathematically equivalent*, though not *identical*. Otherwise we are not warranted in interchanging them.

First of all, then: What operations determine I? As yet, I have not seen them mentioned. One thinks, of course, of some such procedure as this: the introspectively observed sensation whose intensity I is unknown and required is compared with one whose intensity I' is known; and the ratio I/I' is obtained by direct observation. But the psychophysicists from Fechner to the present cry as with one voice that no such operation can be performed; that introspection yields no measure of the intensity of sensation. But if no feasible operation can be specified, the result of which gives a direct measure of the intensity I of introspectional sensation, the question whether I is equivalent to γ becomes a question whether an undefined concept is equivalent to a defined concept; or, in other words, whether a non-existent, impossible operation is equivalent to an actual operation. Busy people usually do not entertain such a question very long after they have detected its character.

We have thus dismissed the question whether I is equivalent to γ, but it remains to show what properties belong to γ considered as a formal entity. It will then be proper to ask whether introspectional experience has yielded data which

fit the form. We are told that the variable values of γ constitute a *scale of measurement* of something that may be called [Fechnerian] sensational intensity. To evaluate this proposition, we should recall a few facts about scales of measurement in general.

On the principle that a magnitude, of any kind whatever, may be expressed by a number, and represented by a length, one may use the following procedure to devise a means of representation called a 'scale.' Taking a directed line-segment whose origin is X_0, one chooses a series of points which divide it into equal segments, and designates the series of points by an arithmetical series of numbers. The absolute difference between any two numbers in the series now denotes a *length—i.e.*, that of the segment limited by the points that the two numbers designate. Each number taken alone designates in general a point. But the distance of any point, such as X, from the origin X_0, is denoted by the expression $X - X_0$; hence each scale number though presented alone can be made to denote a *length* as well as to designate a *point.* The condition to be satisfied is that the first term in the arithmetical series shall be zero, and be applied to the origin X_0 of the line-segment. Then, since $X_0 = 0$, the numerical value of any length $(X - X_0) = (X - 0) = X$. Our scale-numbers may now be used to denote lengths. If they are also to denote the magnitudes which these lengths represent, the necessary condition is that X_0 shall denote a nullity of magnitude. Thus we have devised an absolute scale.

It is often convenient, however, to employ an arbitrary scale instead of the absolute scale. In such case, one arbitrarily chooses some value X' of X as a reference-point, whose number also expresses a reference-magnitude, and specifies every other value of X in terms of its deviation x from X': thus,

$$X_1 = X' + x_1,$$
$$X_2 = X' + x_2, \cdots$$

and in general,

$$X = X' + x$$
$$x = X - X'$$

and in particular,

$$x' = X' - X' = 0$$

whence x' is often called the zero-point on the arbitrary scale. It does not, however, denote absolute nullity of magnitude. That state is denoted on the arbitrary scale by

$$x_0 - x' = X_0 - X' = 0 - X' = - X'$$

Thus, on the arbitary scale, absolute nullity is expressed by $- X'$ while the arbitrary reference-point x' is called zero; whereas on the absolute scale, absolute nullity is called zero and the arbitrary reference-point is denoted by X'. A number on the arbitrary scale, taken by itself, does not express an absolute magnitude. To convert one of these arbitrary numbers, such as x, into an expression of absolute magnitude, X, one puts

$$X = x + X'$$

or more generally,

$$X - X_0 = x + (X' - X_0).$$

However, the difference between two of these arbitrary scale-numbers denotes a difference between two magnitudes, just as the difference between two numbers of the absolute scale. Thus,

$$x_n - x_m = (x_n + X') - (x_m + X') = X_n - X_m$$

and in particular,

$$x - x' = X + X'.$$

We may now turn to Fechner's scale of so-called intensity of sensation. The various values of γ determined by his *Massformel*

$$\gamma = k \cdot \mathrm{Log}\ (\beta/b) \qquad [1]$$

form an array of pure, or abstract, numbers. They are perfectly defined by the operations indicated in the *Massformel*, but as yet they are otherwise meaningless. It has been urged that they have been made to denote a system of empirical magnitudes, called the intensities of sensation, magnitudes which do not owe their existence to the physical,

social, and mathematical operations prescribed by the *Massformel* [1]. However, if we examine the formal characteristics of these alleged magnitudes, we shall find that they do not agree with anything that has yet been empirically found. Here are some of them.

(*a*) *The arbitrary zero-point on the Fechnerian intensity-scale.* In the defining equation [1] put $\beta = b$, noting the meaning of b in terms of its defining operations (*cf.* p. 115 above). This makes the ratio β/b, which we shall designate by the symbol $(\beta/b)'$, equal to unity. From [1],

$$\gamma' = k \cdot \text{Log}\,(\beta/b)' = k \cdot \text{Log}\,1 = k \cdot 0 = 0 \qquad [2]$$

Fechner designated γ' as the null-point on his scale, and asserted that it represented the intensity which a sensation must attain in order to be *sensed* by its owner. If $\beta > b$, $\gamma > \gamma'$, and according to this rule of interpretation, the expression $(\gamma - \gamma')$ represents the excess of intensity which the sensation has over the intensity which it must have in order to be 'sensed.' If $\beta < b$, $\gamma < \gamma'$, and in terms of Fechner's arbitrary rule of renaming, the absolute value of the now negative expression $(\gamma - \gamma')$ represents the deficit of intensity which must be made up before the sensation can be 'sensed.' Thus, negative values of γ were said to represent degrees of 'unconsciousness.' I have already emphasized (*4*, 263–275) that this interpretation is irrelevant to the defining operations; that in terms of the operations peculiar to Fechner's preferred form of the method of right and wrong cases, the proposition $\beta = b$, or that $(\beta/b) = 1$, means simply that the magnitude β of the stimulus is sufficient to produce just 84% of answers classified as 'right'; that the proposition $\beta > b$, or that $(\beta/b) > 1$, means that the magnitude of the stimulus is more than sufficient to produce 84% of 'right' answers, or that the stimulus produced more than 84% of such answers; while the proposition $\beta \leqslant b$, or that $(\beta/b) < 1$, means that the magnitude of the stimulus is insufficient to produce 84% of 'right' answers, or that the stimulus produced fewer than 84% of such answers. The defining operations have nothing to do with degrees of introspectively observed

intensities, or with introspectional consciousness. It is therefore evident to us, as it was to Fechner, that $\gamma' = 0$ on the arbitrary scale only; it does not represent absolute nullity of Fechnerian sensational intensity, nor yet the smallest intensity which a Fechnerian sensation can have. What point γ_0 on the arbitrary scale does represent absolute nullity?

(*b*) *Absolute nullity as represented on the Fechnerian intensity-scale.* It is evident from inspection of the defining equation [1] that the smallest value γ_0 which γ can attain is determined by the smallest value that the ratio β/b can have. Since both β and b are expressions of kinetic Energy or of Work, neither can be negative. Hence, the smallest value $(\beta/b)_0$ which the ratio β/b can assume is nullity. This value would be attained in case β were reduced to zero, or if b should become infinite.[3] Substitution in the defining equation [1] gives

$$\gamma_0 \equiv k \cdot \text{Log}\,(\beta/b)_0 = k \cdot \text{Log}\,0 = k(-\infty) = -\infty. \quad [3]$$

(*c*) *Absolute intensity denoted by Fechner's arbitrary zero.* Fechner's scale being arbitrary, none of its numbers taken singly expresses an absolute magnitude (see p. 117 above). To convert the arbitrary zero γ' into a direct expression Γ' of the absolute magnitude which it denotes, one exhibits Γ' as the deviation of γ' from absolute nullity γ_0. Thus:

$$\Gamma' = \gamma' - \gamma_0 = 0 - (-\infty) = \infty. \quad [4]$$

(*d*) *Absolute intensity denoted by any number on Fechner's scale except* γ_0. The absolute intensity Γ denoted by any arbitrary scale-number γ except γ_0 may be found from first

[3] This is a *modus dicendi*, and must be regarded as being required by a choice of descriptive terms and not by the facts themselves. For example: consider the so-called luminosity-curve, which represents the 'visibility' of a light of constant duration, retinal position and retinal extent, as a function of wave-length. Between some such limits as 400 and 780 millimicrons, this function is finite; it is represented as attaining a zero-value as these limits are transcended. (Such representation, of course, rests on extrapolation beyond the limits of observation.) But 'visibility' by definition is inversely proportional to b. Hence, in saying that the visibility of infra-red or ultra-violet radiation is zero, we say by implication that the intensity at the threshold of effectiveness is infinite.

principles, by substitution of Γ and γ for Γ' and γ' in equation [4], thus:

$$\Gamma = \gamma - \gamma_0 = \gamma - (-\infty) = \gamma + \infty = \infty. \qquad [5]$$

(*e*) *Difference between the magnitudes denoted by any pair of numbers on Fechner's scale excluding the origin* γ_0. From [5] we may take as special instances,

$$\begin{aligned} \Gamma_1 &= \gamma_1 - \gamma_0 = \infty, \\ \Gamma_2 &= \gamma_2 - \gamma_0 = \infty, \\ \Gamma_2 - \Gamma_1 &= \infty - \infty, \end{aligned} \qquad [6]$$

which is indeterminate. Likewise, from [5],

$$\begin{aligned} \gamma_1 &= \Gamma_1 + \gamma_0 = \infty + (-\infty) = \infty - \infty, \\ \gamma_2 &= \Gamma_2 + \gamma_0 = \infty + (-\infty) = \infty - \infty, \\ \gamma_2 - \gamma_1 &= (\infty - \infty) - (\infty - \infty), \end{aligned} \qquad [7]$$

which is indeterminate.

(*f*) *Magnitude of that portion of the absolute intensity which, according to Fechner, is 'sensed.'* This is given by the expression $(\gamma - \gamma')$, in which $\gamma > \gamma'$. Substitution of γ and γ' for γ_2 and γ_1 respectively, in the reasoning which yields equation [7] gives

$$\gamma - \gamma' = (\infty - \infty) - (\infty - \infty), \qquad [8]$$

which is also indeterminate.

(*g*) *Constancy of the absolute intensity denoted by the arbitrary zero.* Suppose two separate experiments yield b' and b'' respectively as the stimulus-magnitudes which yield (say) 84% of responses classified as 'correct.' In the two experiments present $\beta' = b'$ and $\beta'' = b''$ respectively. Then $\beta'/b' = \beta''/b'' = 1$. The corresponding values of γ are γ' and γ'', both of which are represented by the arbitrary zero on Fechner's scale. Do the two arbitrary zeros denote the same absolute intensity? Substitution of γ'' for γ' in equation [4] and of γ' for γ in equation [8] yields

$$\gamma'' - \gamma' = (\infty - \infty) - (\infty - \infty), \qquad [9]$$

which is indeterminate.

While the *Massformel* defines a system of pure numbers as products of a constant into the logarithms of positive ratios, it also defines the properties of the (abstract) magnitudes which these numbers denote. These properties are set forth in equations [2] to [9] inclusive. They correspond to nothing which has yet been exhibited as a datum of human observation.

Summary

1. Fechner's concept of sensational intensity has been examined by the criteria of the operational method of definition. It was found to be the result of three *mathematical* operations made on two *physical* measurements, one of which was selected for treatment by a *social* criterion. The meaning of the concept is this defining procedure.

2. It has been contended by others that while this concept is not identical, in the operational sense, with introspectional intensity of sensation, yet the two concepts are interchangeable in certain relationships, and in such a sense are equivalent. It was here replied that since no feasible operations have been specified which define introspectional intensity of sensation, the latter concept is thus far undefined, and the question of its equivalence to Fechner's concept is meaningless.

3. An analysis has been made of the γ-system of numbers defined by Fechner's *Massformel* and discussed by him as if they denoted magnitudes on a scale of sensational intensity. This number-system was shown to have the following properties: its first member denotes absolute nullity; every other member denotes infinity. All comparisons of such numbers yield indeterminate quantities; therefore all comparisons of the magnitudes which may be denoted by them are meaningless. This assertion applies whether the magnitudes are expressed as deviations from absolute nullity, or from the arbitrary zero on Fechner's scale. The constancy of Fechner's arbitrary zero was also shown to be indeterminate.

Conclusion

If anyone can find introspectional data the intensities of which can be observed to have the properties denoted by

Fechner's system of scale-numbers, he is to be congratulated on having participated in a miracle. But if the intensities of his data of introspection are observed to be finite, if he can even make distinctions among their intensities, the latter are not denoted by members of this number-system. In other words, under such conditions, his sensations are not Fechnerian.

References

1. Bridgman, P. W. The logic of modern physics, New York, Macmillan, 1927.
2. Fechner, G. T. Elemente der Psychophysik, 2 v., Leipzig, Breitkopf und Hartel, 1860.
3. Fechner, G. T. In Sachen der Psychophysik. Leipzig, Breitkopf und Hartel, 1877. *Vide* especially Ch. 10.
4. Johnson, H. M. Did Fechner measure '*introspectional*' sensations? Psychol. Rev., 1929, 36, 257–284.
5. Mach, E. Popular scientific lectures (T. J. McCormack, Tr.), Chicago, Open Court Publishing Co., 1910. *Vide* especially the lecture entitled 'The economical nature of physical inquiry' for an exemplification of operational definition.

[MS. received August 31, 1929]

THE PRESENT STATUS OF EXPERIMENTAL STUDIES ON THE RELATIONSHIP OF FEELING TO MEMORY

BY H. MELTZER

St. Louis Psychiatric Clinic and Washington University

There is no denying the importance of the Freudian interpretation of forgetting. It is a "conception which forms the foundation stone on which the whole structure of psychoanalysis rests." [1] And though not generally acceptable to experimental psychologists, it is influencing, or is at least closely related to, the thinking of many students from many special fields of knowledge. Clinical psychologists [2] and psychiatrists (37) [3] see in it the explanation for many unfortunate peculiarities of behavior. There is an obvious relationship between this concept and the law of effect of educational psychologists.[4] Sociologists like Ogburn [5] see in this tendency to forget the unpleasant a possible explanation for our conservatism and our cultural inertia. And Bertrand Russell, in his writings,[6] makes frequent use of this as well as other concepts from Freudian psychopathology. So much for applications and possible implications; what are the known facts? What is the experimental evidence? What facts are known concerning human remembering and forgetting of pleasant and unpleasant experiences? How were these facts ascertained? How justified were the assumptions

[1] T. W. Mitchell, Problems in psychopathology, Harcourt Brace, 1927, p. 29.

[2] See, *e.g.*, Morgan's The psychology of the unadjusted school child, Macmillan, 1925, p. 75.

[3] Numbers in parentheses refer to bibliography accompanying this article.

[4] *Cf.* E. L. Thorndike, Educational psychology, vol. III., or Gates' Psychology for students of education. Sandiford, in a later work (Educational psychology, Longmans, 1928, p. 248) says that, to the non-Freudian, the phenomena of repression seem to be due to the non-exercise of the unpleasant and adds that most cases cited in psychoanalytic literature appear to be explained more rationally by the laws of exercise and effect.

[5] Social change, B. W. Huebsch, 1922, p. 186.

[6] See, *e.g.*, his Sceptical essays, Norton, 1928.

made by the investigators, how reliable were the methods used, and how warranted the conclusions drawn? How revealing were these studies? Why—or why not? And finally, what is the total accumulative gain from the point of view of research? . . . In the present study an attempt will be made to evaluate the present status of this problem, and, with the foregoing questions in mind, analyze the influences which are retarding and advancing this aspect of affective psychology.

Experimental Investigations in Chronological Order

The present survey of the literature is limited to experimental investigations. For a general discussion of the hedonistic theories of motivation the reader is referred to Troland's recent book (**42**). The survey includes all studies summarized and discussed—usually reinterpreted—by Wohlgemuth (1923) in what was purposed to be an exhaustive review of the experimental literature on the general problem of the influence of feeling on memory. From his list he omitted investigations concerned with such narrower aspects as the rôle of feeling tone in reproduction and association time. The references cited and discussed by Laird (**20**), Griffits (**13**) and Stone (**36**), in their historical approaches, are, of course, all included. And the present survey includes finally all other relevant experimental studies found by the present writer after examining all available sources.

From the studies listed in the *Psychological Index* under the caption: Affection, Hedonic Tone (Pleasantness and Unpleasantness), were omitted studies on color preferences, philosophical discussions of such problems as the moral synthesis between feeling and idea, and introspective studies like that of Nafe [7] which deal with refined abstractions rather than relationship between feeling and memory. The list is, then, fairly inclusive and approaches being exhaustive.

The investigations, beginning with the earliest found in literature, will be reported in chronological order (correct by years, not months; issues of monthly and quarterly journals

[7] *Amer. J. Psychol.*, 1924, 35, 507 *ff.*

LIST OF EXPERIMENTAL STUDIES

Investigator and Date	Subjects	Method	Conc. Results	Remarks
1. Colgrove, 1898.	Large number school children.	Questionnaire.	+ Pleasure predominant.	Question poorly stated and susceptible to misinterpretation. Method highly unreliable.
2. Kowalewski, 1904.	270 school children, ages 10–13.	Questionnaire.	+ Percentage in favor of P: 69.4 for boys; 67.8 for girls.	Method unreliable though question somewhat more clearly stated than first study.
3. Gordon, 1905.	7 subjects; 6 Ph.D.'s, 1 student; 5 M., 2 F.	Recognition of colored figures and protocols after 3 weeks.	– No essential difference between U, P and I.	First more objective approach. Insufficient number of subjects.
4. Kowalewski, 1908.	105 elementary school boys, 9–15.	Recall 10 days after writing out P and U experiences of previous day.	+ Primary memory optimists [8] 69%; secondary 61.4%.	Did not correct for 'new' experiences—those given second time and not the first.
5. Peters, 1911.	8 subjects.	Discrete association [9] (948 experiments).	+ 80% experiences affectively toned. Of these were: 65% P; 30% U; 5% mixed feeling tone.	Insufficient number of cases.
6. Henderson, 1911.	10 college students.	Relate 100 incidents in life and grade as P, U and I.	–? Recalled: 55.1% agreeable; 11.8% indiff.; 33.1% disagr.	Insufficient no. of cases. Results and conclusion not in agreement. Argues that results do not show that we forget disagreeable, but majority of exp. of life not disagreeable.
7. Birnbaum, 1912.	5 hospital attendants.	Reproduction or association time.	– No difference in influence of feeling tones.	Insufficient number of cases.

[8] By a primary optimist Kowalewski meant an individual who on first attempt to recall gave more pleasant than unpleasant experiences; a secondary optimist, then, is one who gave more pleasant experiences on second recall.

[9] Subject was asked to respond to given stimulus word by recalling previous actual experiences, and grade feeling tone of original experience as well as memory of it.

LIST OF EXPERIMENTAL STUDIES—*Continued*

Investigator and Date	Subjects	Method	Conc. Results	Remarks
8. Tait, 1913.	9 subjects.	Reproduction of words and recognition of colors.	+ Difference of 6% in favor of P. P also has some influence on others. I remembered worst.	Assume equal number of pleasant and unpleasant experiences.
9. Peters and Nemecek, 1914.	118 students, 14–24, and 28 children, 10–11.	Discrete association method.	+ Pleasant 50.2%, unpleasant 45.6%, mixed 4.2%.	Assumed same no. of pleasant and unpleasant experiences to recall.
10. Myers, 1915.	232 persons from normal, high and college.	List 100 random objects and events and classify according to likes and dislikes.	+ Recall thing best liked.	Method and results similar to (6) Henderson, but investigator's conclusion more in accord with results.
11. Tolman, 1917.	6 subjects; 3 psych. students, 3 naive.	Memory of list of pleasant, unpleasant and indifferent words.	+ 11.5 repetitions for P. 12.7 for I. 13.5 for U.	Feeling tone of list of words predetermined, not ascertained by verbal report of subjects. Results differ from Tait's in placing memorability of I above U.
12. Baxter, Yamada and Washburn, 1917.	69 college women.	Directed recall: free association method followed by classification of exp. as physical or mental.	+ Time recall for unpleasant is longer. Physical unpleasant more readily recalled.	Conditions too narrowly restricted to elicit responses characteristic of everyday life.
13. Tolman and Johnson, 1918.	35; 18 males, 17 females.	Average association reaction time for P, U, and I words.	+ Time longer for unpleasant than pleasant. Order of memorability is P, U, I.	Feeling tone of words not ascertained by verbal report of subject; words chosen by investigator and same list used for all subjects.

LIST OF EXPERIMENTAL STUDIES—*Continued*

Investigator and Date	Subjects	Method	Conc. Results	Remarks
14. Morgan, Mull and Washburn, 1919.	97 college women.	Reaction time to given words followed by grading stimulus P or U. Score being no. of P and U words suggested.	+ More pleasant words suggested. Positive correspondence with judgments of friends on temperament.	Stimulus too controlled and restricted. Inference indirect. Not primarily designed to throw light on general problem of rôle of feeling in memory.
15. Griffits, 1920.	36 college students.	4 experiments. a. Discrete association followed after 1/2 hour by recall. b. Visual presentation. c. Different list words. d. Continuous ass'n method.	− More pleasant than unpleasant associations but no evidence for repression.	Stimulus words used divided equally between those which *might* be expected to call up pleasant or unpleasant associations. They *might not* for some of subjects. Situation not lifelike.
16. Smith, 1921.	50 on galvanic reflex, 49 on reaction time and 22 on Jung Test.	Psychogalvanic reflex, reaction time, Jung Reproduction Test.	+ Positive tone facilitates, negative inhibits memory.	Not easily comprehensible because relation of positive and negative tone to pleasantness and unpleasantness is a sheer assumption for which there is no evidence.
17. Fox, 1923.	24 men, ages 22–38½.	Learning and retention of sonnets followed by subjects indicating their preferences.	+ Preferred sonnets better learned and better retained.	Preferences made after rather than before learning. Facility in learning might be cause for preference.
18. Laird, 1923.	62 beginning students in psychology.	Write 10 names of friends and check those liked and disliked.	− Recall of pleasant dependent upon temperament. No evidence for repression.	Experiment not primarily designed to throw light on more general problem.

LIST OF EXPERIMENTAL STUDIES—*Continued*

Investigator and Date	Subjects	Method	Conc. Results	Remarks
19. Wohlgemuth, 1923.	687 school children, ages 11–16.	Recall of remembered personal experiences. Time gap from 1st to 2d recall 10 days or fortnight.	− 40.1% pleasant and 39.8% unpleasant forgotten. Average for each child: pleasant experiences forgotten, 38.2%; unpleasant, 35.3%.	Profited by errors of previous investigators but too interested in disproving Freud. Adult subjects probably more suited for this problem. Analysis of individual differences not made.
20. Stone, 1925.	12 sophomores; 10 males, 2 females.	Name likes and dislikes. Photographs of persons never seen with name allotted shown month later.	+ Pleasant impressions remembered better.	Time exposure not accurately kept. Physiological expressions in specific emotions also reported.
21. Anderson and Bolton, 1925.	80 (mostly students in elem. psychology).	Reading times of words; recall and recognition of odors with nonsense names attached.	− a. Odor experiment: % Recall Recog. P..... 30 71 U..... 29 73 I..... 27 62 b. Average reading time; in sigma: P, 347, U, 351; words recalled: P, 7; U, 5.1.	Situations not sufficiently life-like
22. Gordon, 1925.	200; 3 instructors, 197 students; 11 men, 189 women.	Recollection of 10 pleasant and unpleasant odors in liquid form.	− % recalled: P, 49; U, 50; I, 56. Correlation of −.07 between memory value and affective value.	Technique objective and analysis of problem good. Too large a proportion of subjects of one sex. Stimuli not life-like.

LIST OF EXPERIMENTAL STUDIES—*Continued*

Investigator and Date	Subjects	Method	Conc. Results	Remarks
23. Gordon, 1925.	750 students.	Analysis of description of earliest memories.	– Averages for: males, 3.58 P, 3.5 U; females, 3.79 P, 3.61 U. % memories for: males, 25 P, 53.2 U; females, 35.2 P, 46 U.	Designed to note active process of repression, not measure forgetting of experiences previously known to be remembered.
24. Flugel, 1925.	9 subjects.	Record affective life for 30 days.	+ Predominance of pleasure over unpleasure in duration and intensity but degree of predominance varying from one subject to another.	Refers to duration rather than number of experiences. Reinterpreted by Wohlgemuth to give opposite conclusion.
25. Kenneth, 1927.	29 males; 34 females.	Affective and associative reactions to 26 osmyls.	+ In many cases olfactory stimuli aroused forgotten experiences.	Use of olfactory stimuli in psychoanalysis suggested by investigator. Hollingworth would doubtless explain as redintegrative process. Not essentially an exp. in relation of feeling and memory.
26. Chaney and Lauer, 1929.	20 girls, ages 18–21.	Recall and retention. Test given after lunch 5 consecutive days.	– % recalled: 72.2 P; 71.4 I; 72.9 U.	Same list of words used for all subjects and feeling tone of words evaluated but not individually by subject. Materials used comparatively meaningless.

are not temporally comparable enough to warrant a finer chronological listing) and presented in tabular form. The extent of information reported concerning each study is indicated by the captions. These are practically all self-explanatory. Only two call for comment. Under 'Conclusion,' plus (+) and minus (−) signs are used. Plus (+) indicates that results, or interpretation of same by the investigator, favor the conclusion that pleasant feeling tone facilitates memory and that we tend to inhibit the unpleasant. Minus (−) indicates that results or interpretation made by investigator disprove or oppose such a conclusion. The other item tabulated on which comment should be made is 'Remarks.' This would have been probably more accurately captioned 'Errors,' for the comments therein included are usually descriptions of shortcomings of method used or unwarranted assumptions or interpretations made. It should be remembered that this does not imply that studies included are characterized only by shortcomings. P, U, and I are symbols for pleasant or pleasantness, unpleasant or unpleasantness, and indifferent, respectively.

In the foregoing studies there is no absence of the 'confusion and contradiction' for which Young (**52**) tells us affective psychology is notorious. The range of disagreement is as wide as it possibly could be. To Tait, who may be taken as a representative of one extreme, "the obvious conclusion is that pleasant words are better remembered and exert an influence on others" (**39**). Myers (**24**), Tolman and Johnson (**41**), and Washburn and her students (**43–47**), using different methods, also draw similar conclusions. Griffits (**13**) steers as near a middle course as any investigator. Wohlgemuth, who represents the other extreme, after reviewing previous investigations as well as reporting his own, comes to the conclusion that "there is no difference whatever between the two feeling tones, pleasure and unpleasure, in their influence upon memory" (**50**). Kate Gordon, back in 1905, in reporting what may be considered the first more objective approach to this problem, came to the conclusion that "Das Wiederkennen des früher gesehenen Figuren ist

gleich gut für die gefälligen, missfälligen und gleichgültigen" (12). Her interest in this problem has been persistent and in 1925 (10) and again in 1928 (11) we find her reporting studies using other materials and other methods but the conclusion is in effect the same—there is no difference, and there is no evidence for repression!

Even among the investigators who agree on the general conclusion that pleasantness facilitates remembering there is no agreement on the order of sequence for unpleasant and indifferent stimuli. Tolman (40) places memorability of indifferent stimuli above the unpleasant whereas Tait (39) places the memorability of unpleasant above that of indifferent. Griffits (13) as well as Anderson and Bolton (1) find no difference between pleasant and unpleasant but conclude that affectively toned stimuli—pleasant or unpleasant—have a factor of advantage over indifferent stimuli.

The nature of the shortcomings of these studies may briefly be classified as follows: Insufficient number of subjects, 4; entire method unreliable (questionnaire studies), 2; aspect of methodology faulty, 6; conditions too narrowly restricted and not sufficiently life-like, 5; unwarranted assumptions tacitly made, 4; and forced or unwarranted interpretation, 3.

The association method was used by many of the investigators. And what Laird's experimental study [10] reveals about the invalidity of the association test to a practical situation such as is presented in advertising holds equally of its use for the study of feeling reactions. In the words of the investigators who attempted the use of the psychogalvanic reflex for dream analysis, "The word-association method has produced interesting results, but seldom helps one to penetrate into the depths of personality. It isn't the dynamite charge of the miner, but the blunderbuss filled with spreading shot. Most of it misses its intended victim. A little penetrates, but not far. And even of these pellets which make him wince, some have only too obviously struck him on parts which are tender to most people." [11] A similar criticism can

[10] A critique of the association test as applied to advertising, J. Exper. Psychol., 1923, 6, 357–365.

[11] A. G. Ikin, T. H. Pear and R. H. Thouless, The psychogalvanic phenomenon in dream analysis, *Brit. J. Psychol.*, 1924–25, 15, 23–43.

be given of the use of odors. The recollection of pleasant and unpleasant odors is not the same as the recollection of pleasant and unpleasant experiences in everyday life. With the exception of Kowalewski and Wohlgemuth, who studied actual personal experiences of school children, none of the investigators studied the reactions of the organism-as-a-whole. Instead, the methods used practically imply that a personality is a summation of discrete responses to discrete stimuli. And, while such an approach is invaluable to the structuralist in ascertaining how the mind is composed, it is not applicable to the study of feeling reactions in everyday life. Instead of analyzing the total emotional situation including the social setting, the investigators, by and large, have studied their individuals in artificial laboratory situations. And it is this fact, more perhaps than any other, that explains the lack of progress made in the investigation of this problem.

Of the tacit assumptions made which invalidate the results reported by these investigators, the one most frequently found is that we experience the same number of pleasant and unpleasant activities and hence that we can correctly infer the relative potencies of pleasantness and unpleasantness by studying the ratio of unpleasant to pleasant experiences in recall. Griffits and Wohlgemuth are the only investigators who express their awareness of this unwarranted assumption and its implications.

Another factor which has retarded advancement is the truncated general approach of the investigators. This can often be traced to psychoanalytic influence. Now there is no denying our indebtedness to Freud for insights into human nature which experimental psychologists, working with more normal people under controlled, but also narrowly restricted, conditions of the laboratory, did not perceive and perhaps would not have perceived for many years, if ever. But there is also no denying that Freudian thinking is compartmentalized thinking, and that his theory of forgetting is, characteristically enough, a large generalization. And, 'too large a generalization,' as Whitehead informs us, 'leads to mere barrenness.' [12] In psychoanalytic literature things happen

[12] Science and the modern world, Macmillan, 1928, p. 43.

'invariably' and 'universally.' Classification, not measurement, is the general procedure. One is either 'this' or 'that.' Things 'are' or 'are not.' The result is that in the study of this aspect of the problem of memory we find ourselves in a stage of development not much further advanced, as far as general approach is concerned, than was the status of the general study of memory before the classical study of Ebbinghaus. Early studies of memory revealed little that was significant because, as Gardner Murphy (**25**) points out, psychologists "were thinking in terms of learning versus forgetting, making a sharp line between what was learned and what was not learned, between what was forgotten and what was not forgotten. They were not yet thinking in quantitative terms; they took no account of degrees of learning and degrees of forgetting."

Investigators of the influence of feeling on memory—an aspect of the problem of memory neglected by the followers of Ebbinghaus, who have limited their studies to work with nonsense syllables, multiplication tables and other relatively meaningless materials rather than personal experiences [13]—are making precisely the same error. They are attempting to discern whether Freud is right or wrong; whether unpleasant feeling tone leads to forgetting or does not; whether human beings repress the unpleasant or do not repress the unpleasant; whether the 'law of effect' is right or wrong, instead of measuring individual differences in forgetting pleasant and unpleasant experiences. Some of the investigators have become 'enlightened' to see forgetting as an active process and repression as the mental mechanism explaining it. Others, like Wohlgemuth, have become entirely 'disillusioned' and therefore seek to controvert the dictum "that unpleasant experiences were more easily forgotten than pleasant ones" [14] along with other psychoanalytic dicta. We thus find protagonists and antagonists all remaining on that midway stage of development—classification. In Aristotelian logic, classification rather than measurement

[13] See chapter on 'How we forget' in (**27**) for distinction between materials which have 'logical relevance' and 'affective relevance.'

[14] *Op. cit.* (**50**), p. 243.

was the means of discovery. And, as Whitehead tells us, "The popularity of Aristotelian logic retarded the advance of physical sciences throughout the Middle Ages." [15] The popularity of the 'new psychology' with its easily classifiable labels may serve the science of psychology as Aristotelian logic served the physical sciences.

Investigators of this particular problem, many of whom are familiar enough with psychological history, have permitted Freudian influences to truncate their thinking and, as a consequence, they have neglected the contributions concerning individual differences that have come to us by way of Galton, Cattell, Thorndike and their followers. The result has been that they have searched for facts which prove or disprove a conclusion instead of measuring realities. This immediate seeking for explanatory principles correlates with the attitude of mind in which association is interpreted as necessarily implying causation. The fact that in his experiment he found pleasantness associated with better remembering, Tait, for example, interpreted as proving that pleasantness is the cause of improved remembering. The fact that in his experiment, Wohlgemuth found no such relationship, he interprets as proof that feeling does not influence memory. To one approaching the problem without preconceived notions —Freudian or anti-Freudian, Thorndikian or anti-Thorndikian—such interpretation seems neither logical nor factual. Association, even close association, does not necessarily imply causation and the facts reported by the investigators do not prove or disprove such causation. As likely an explanation, as it appears to the writer, is that the conditions and traits which influence learning also influence feeling reactions.

The argument here is not against analysis, for the writer agrees with Professor Wheeler that "Regarding things as wholes, no matter how much esthetic satisfaction or mental repose one may derive from their contemplation, is not scientific explanation." [16] On the contrary, this is an argument for analysis—but the analysis of the problem, the statement of which does not make the aforementioned pre-

[15] *Op. cit.*, p. 43.

[16] Present tendencies in biological theory, *Scient. Mo.*, Feb. 1929, p. 97.

suppositions. This implies the need for the study of, first, not the influence of feeling on memory, but the study of such unitary responses as the memory of pleasant experiences, the forgetting of pleasant experiences, the memory, or the forgetting, of unpleasant experiences. These unitary responses [17] are not unanalyzable wholes, but are identifiable and measurable forms of realities.

In spite of all the shortcomings in methods used and disagreement in interpretations made, advancement, however, can be observed. From Colgrove's (6) tabulation of children's answers to the question, "Do you recall pleasant or unpleasant experiences better?" to Wohlgemuth's (50) analyses of children's reports of pleasant and unpleasant experiences, taken once on a day following a vacation and taken again about a forthnight later, or to the psychogalvanic reflex method used by Smith (35) with all its incomprehensiveness, represents real progress in methodology. Similar signs of advancement may be observed in 'things which go without saying'—the tacit assumptions made by investigators which further study proves unwarranted; the general approach to the problem; the treatment of the data and the interpretation of results. Though not as large as we should like it to be, there has been some accumulative gain

Conclusion

A critical survey of the experimental literature on this problem clearly suggests the need for a redirection of emphasis from classification and interpretation in terms of preconceived theories and biases to the measurement of realities. In such a redirection it would be well to bear in mind the specific shortcomings of the previous investigations reported in this study, as well as the general features of psychoanalysis which make for retardation or advancement. The distinctive feature of psychoanalysis which has made for progress in our understanding of feeling reactions is its emphasis on the entire emotional situation including the social conditions. The feature which makes for retardation is its tendency to facilitate

[17] In the Woodworth sense, similar in meaning to 'configurational responses.'

compartmentalized thinking. With this tendency is associated the tendency to interpret association as necessarily implying causation. The question for which the investigators have sought an answer has been: Does feeling tone —pleasantness or unpleasantness—*influence* remembering or forgetting? This manner of stating the problem presupposes that feeling conditions memory. As likely an explanation for associations found between these two processes is that much of what conditions learning also conditions our feeling reactions. To make this problem more susceptible to investigation and verification the question would be better restated in terms which make neither of these assumptions and which treats the memory or the forgetting of a pleasant or unpleasant experience as unitary responses. Such redirection may not immediately reveal explanatory principles, but it will yield reliable information concerning such significant effects as the extent to which individuals differ in experiencing pleasant and unpleasant activities in everyday life, the extent to which individuals differ in forgetting pleasant and unpleasant experiences, and the extent to which such experiencing and forgetting correlates with other significant measurable effects. It is the opinion of the present writer that, unless such a redirection of approach takes place, these questions will remain unanswered and 'confusion and contradiction' will continue to characterize the status of our knowledge concerning this aspect of affective psychology.

BIBLIOGRAPHY

1. ANDERSON, A. C., AND BOLTON, F. J. The inhibition of the unpleasant, *J. Abn. & Soc. Psychol.*, 1925, **20**, 300–302.
2. BALLARD, D. B. Obliviscence and reminiscence, *Brit. J. Psychol., Monog. Supp.*, vol. I., 1915.
3. BEEBE-CENTER, J. G. Law of affective equilibrium, *Amer. J. Psychol.*, 1929, **41**, 1, 54–69.
4. BIRNBAUM, K. Ueber den Einfluss von Gefühlsfaktoren auf die Assoziationer, *Monatschr. f. Psychiat. u. Neurol.*, 1912, **32**, 95–123.
5. CHANEY, R. M., AND LAUER, A. R. The influence of affective tone on learning and retention, *J. Educ. Psychol.*, 1929, **20**, 287–291.
6. COLGROVE, F. W. Memory, London, 1901.
7. CROSLAND, H. R. A qualitative analysis of the process of forgetting, PSYCHOL. MONOG., 1921, **29**, 130.
8. FOX, C. The influence of subjective preference on memory, *Brit. J. Psychol.*, 1922–23, **13**, 398–405.

9. FLÜGEL, J. C. A quantitative study of feeling and emotion in everyday life, *Brit. J. Psychol.*, 1925, **15**, 318–355.
10. GORDON, K. Recollection of pleasant and unpleasant odors, J. EXPER. PSYCHOL., 1925, **8**, 225–239.
11. GORDON, K. A study of early memories, *J. Delinq.*, 1928, **12**, 129–132.
12. GORDON, K. Ueber das Gedächtnis für Affectiv bestimmte Eindrücke, *Arch. f. d. ges. Psychol.*, 1905, **4**, 437–458.
13. GRIFFITS, C. H. Results of some experiments on affection, distribution of associations and recall. J. EXPER. PSYCHOL., 1920, **3**, 447–464.
14. HACKER, F. Systematische Traumbeobachtungen mit besonderer Berücksichtigung der Gedanken, *Arch. f. d. ges. Psychol.*, 1911, **21**, 1–131.
15. HENDERSON, E. N. Do we forget the disagreeable?, *J. Phil., Psychol. & Sci. Meth.*, 1911, **8**, 432–438.
16. HOLLINGWORTH, H. The obliviscence of the disagreeable, *J. Phil., Psychol. & Sci. Meth.*, 1910, **7**, 709–714.
17. KENNETH, J. H. An experimental study of affects and associations due to certain odors, PSYCHOL. MONOG., 1927, **37**, pp. 64.
18. KOWALEWSKI, A. Schopenhauer und seine Weltanschauung, Berlin, 1908.
19. KOWALEWSKI, A. Studien zur Psychologie des Pessimismus, Wiesbaden, 1904.
20. LAIRD, D. The influence of likes and dislikes on memory as related to personality, J. EXPER. PSYCHOL., 1923, **6**, 294–300.
21. LOBSIEN, M. Das Gedächtnis, 1913, pp. 91.
22. MEUMANN, E. Psychology of learning, New York, London, 1913.
23. MEYER, M. F. Nervous correlate of pleasantness and unpleasantness, PSYCHOL. REV., 1908, **15**, 201–206, 292–322.
24. MYERS, G. C. Affective factors in recall, *J. Phil., Psychol. & Sci. Meth.*, 1915, **12**, 85–92.
25. MURPHY, G. An historical introduction to modern psychology, New York, 1929, Ch. XI.
26. OFFNER, M. Das Gedächtnis, Berlin, 1913, 2d ed., pp. 85.
27. PEAR, W. H. Remembering and forgetting, London, 1922, Chap. IX.
28. PETERS, W. Gefühl und Erinnerung; Beiträge zur Erinnerungsanalyse, *Psychol. Arbeit.*, 1911, **6**, 197–260.
29. PETERS UND NEMECEK. Massenversuche über Erinnerungsassoziationen, *Fortschr. d. Psychol. u. i. Anwend.*, 1914, **2**, 226–245.
30. RANSCHBURG, P. Ueber Hemmung gleichzeitiger Reizwirkungen, *Zsch. f. Psychol.*, 1902, **30**, 398.
31. RIVERS, W. H. R. Instinct and the unconscious, Cambridge, 1920, App. I.
32. RUCKMICK, C. A. The psychology of pleasantness, PSYCHOL. REV., 1925, **32**, 362–383.
33. SKAGGS, E. B. The concept of retroactive inhibition, PSYCHOL. REV., 1926, **33**, 237–244.
34. SMITH, W. W. The measurement of emotions, New York, 1922.
35. SMITH, W. W. Experiments on memory and affective tone, *Brit. J. Psychol.*, 1921, **11**, 236–250.
36. STONE, A. R. The reaction of memory to affective states, *Amer. J. Psychol.*, 1925, **36**, 112–123.
37. Symposium on Rôle of repression in forgetting. Participants: PEAR, T. H.; WOLF, A. A.; MITCHELL, T. W.; and LOVEDAY, T. *Brit. J. Psychol.*, 1914–15, **7**, 139–165.

38. Szymanski, J. S. Gefühl und Erkennen, Berlin, 1926.
39. Tait, W. D. Effect of psychophysical attitudes on memory, *J. Abn. & Soc. Psychol.*, 1913–14, **8**, 10–38.
40. Tolman, E. C. Retroactive inhibition as affected by conditions of learning, Psychol. Monog., 1918, **107**, pp. 22 *ff.*
41. Tolman, E. C., and Johnson, I. A note on association time and feeling, *Amer. J. Psychol.*, 1918, **29**, 187–195.
42. Troland, J. T. Fundamentals of human motivation, New York, 1928, Chs. 16, 17.
43–47. Vassar Studies. Washburn, M. F., and students.
43. Baxter, M. F., Yamada, K., and Washburn, M. F. Directed recall of pleasant and unpleasant experiences, *Amer. J. Psychol.*, 1917, **28**, 155–57.
44. Graves, K., Heath, E., and Washburn, M. F. Directed egocentric responses, *Amer. J. Psychol.*, 1919, **30**, 300–302.
45. Morgan, E., Mull, H. K., and Washburn, M. F. An attempt to test moods or temperaments of cheerfulness and depression by directed recall of emotionally toned experiences, *Amer. J. Psychol.*, 1919, **30**, 302–304.
46. Washburn, M. F., Giang, F., Ives, M., and Pollock, M. Memory revivals of emotions as a test of emotional and phlegmatic temperaments, *Amer. J. Psychol.*, 1925, **36**, 456–459.
47. Washburn, M. F., Harding, L., Simmons, H., and Tomlinson, D. Further experiments on directed recall as a test of cheerful and depressed temperaments, *Amer. J. Psychol.*, 1925, **36**, 454–456.
48. Washburn, M. F. Feeling and emotion, Psychol. Bull., 1927, **24**, 573–595.
49. Wittenberg Symposium on Feelings and Emotions. Clark Univ. Press, Worcester, Mass., 1928. Especially:
(a) Bühler, K. Displeasure and pleasure in relation to activity. XIV.
(b) Langfeld, H. The rôle of feeling and emotion in æsthetics. XXX.
(c) Dunlap, K. Emotion as a dynamic background. XI.
(d) Weiss, A. P. Feeling and emotion as forms of behavior. XIII.
(e) Jorgensen, C. A theory of the elements in the emotions. XXVI.
50. Wohlgemuth, A. A critical examination of psychoanalysis, London, 1923, pp. 28–39 and 209–217.
51. Wohlgemuth, A. The influence of feeling on memory, *Brit. J. Psychol.*, 1922–23, **13**, 405–429.
52. Young, P. T. Studies in affective psychology, *Amer. J. Psychol.*, 1927, **38**, 157–193.

[MS. received August 26, 1929]

THE RELIABILITY COEFFICIENT AS A PER CENT, WITH APPLICATION TO CORRELATION BETWEEN ABILITIES *

BY ROBERT CHOATE TRYON

National Research Council Fellow, University of California

The reliability coefficient of a variable, X, is a special type of correlation coefficient which indicates the degree to which individuals systematically differ from each other in the trait as measured. The systematic factors at work may be of complex origin, *i.e.*, they may be a composite of independent component systematic factors, but as a compound they are called the *true* ability, X_∞, measured by the test. The unsystematic factors, which never consistently operate in the same direction on the same individual, are occasioned by a multiplicity of chance factors which make the individual differ in performance from his true score. The discrepancy between the true score, X_∞, and the fallible obtained score, X, is a function of these chance factors and is called an error of measurement, E. The purpose of this paper is to show that the reliability coefficient as it stands is a value denoting the *per cent* of the fallible variable, X, produced by X_∞, and further, it will attempt to show that the composition of two correlated variables, X and Y, is readily reducible to per cents by which each is determined (1) by the common factor or ability which occasions the correlation between the traits, (2) by respective independent residual abilities (remaining systematic factors) left over in each variable after eliminating the common factor, and (3) by their respective errors of measurement. The percentage determination of X_∞ and Y_∞ by the common factor and by their respective residual abilities will also be given.

* I wish to express my gratitude to Professor H. E. Garrett and to Dr. J. Yoshioka for their reading of this paper in manuscript.

Proof that the Reliability Coefficient is a Per Cent

In computing a reliability coefficient, one must have two fallible measures, X_1 and X_2, of the same trait or ability. The reason that $X_1 \neq X_2 \neq X_\infty$ is explained by the presence of adulterating errors of measurement in X_1 and X_2. The postulation is that

$$x_1 = x_\infty + e_1, \tag{1}$$

$$x_2 = x_\infty + e_2. \tag{2}$$

The e's are the respective errors in x_1 and x_2, and are uncorrelated with each other or with x_∞. Diagrammatically, the situation is as in Fig. 1. The oval to the left represents

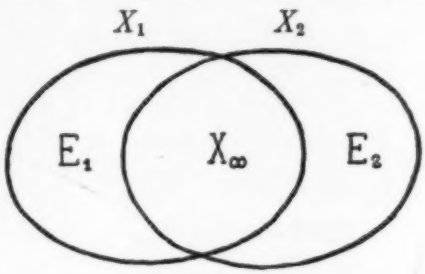

Fig. 1

the fallible variable, X_1, the one to the right, X_2, the common portion is X_∞ and the outer portions are errors of measurement.

In a recent paper,[1] the writer showed that with such a set-up as in Fig. 1, then

$$r^2_{x_\infty x_1} + r^2_{e_1 x_1} = 1, \tag{3}$$

$$r^2_{x_\infty x_2} + r^2_{e_2 x_2} = 1. \tag{4}$$

Equation (3) indicates that $r^2_{x_\infty x_1}$ is a per cent,[2] for where X_1 is considered to have a total composition of *unity* (1.00), then the value, $r^2_{x_\infty x_1}$, is the amount that X_∞ contributes to that whole, in terms of correlation. The value $r^2_{e_1 x_1}$ supplies the

[1] R. C. Tryon, The interpretation of the correlation coefficient. Psychol. Rev., 1929, 36, 419–445.

[2] R. C. Tryon, *op. cit.* See Fig. 1, and proof of formula (1).

remaining per cent of total possible correlation of which X_1 is capable (namely, unity). The amount $r^2_{x_\infty x_1}$ is called the *percentage degree of determination of* X_1 *by* X_∞, and the value, $r^2_{e_1 x_1}$, *the percentage degree of determination of* X_1 *by* E_1, *the errors of measurement.* X_2 has similar percentages as shown in (4).

Likewise, it is true that

$$\frac{\sigma^2_\infty}{\sigma^2_1} + \frac{\sigma^2_{e_1}}{\sigma^2_1} = 1, \tag{5}$$[3]

$$\frac{\sigma^2_\infty}{\sigma^2_2} + \frac{\sigma^2_{e_2}}{\sigma^2_2} = 1. \tag{6}$$

Thus, from (5), the ratio of the variance (σ^2) of X_∞ to the variance of X_1 plus the ratio of the variance of E_1 to that of X_1 equals unity or 100 per cent. The expression $\sigma^2_\infty/\sigma^2_1$ and $\sigma^2_{e_1}/\sigma^2_1$ are called the percentage degrees of determination of X_1 by X_∞ and E_1, respectively, in terms of *total variance of* X_1.

But the two *types* of per cents are equal,[4] that is,

$$r^2_{x_\infty x_1} = \sigma^2_\infty/\sigma^2_1 \quad \text{and} \quad r^2_{x_\infty x_2} = \sigma^2_\infty/\sigma^2_2, \tag{7}$$

$$r^2_{e_1 x_1} = \sigma^2_{e_1}/\sigma^2_1 \quad \text{and} \quad r^2_{e_2 x_2} = \sigma^2_{e_2}/\sigma^2_2. \tag{8}$$

These relationships show that the percentage degree of determination of a composite variable, X_1, by one of its components (such as X_∞ or E_1) may be understood either to be in terms of per cent of total variance of X_1 or in terms of total possible (unity) correlation of X_1 which the given component variable contributes. When the r^2's of (7) and (8) are secured, then the per cents are obtainable.

Now, recall that the definition of the reliability coefficient is the correlation between X_1 and X_2, namely, r_{12}. For this coefficient to be a valid one, the stipulations are[5] that

[3] R. C. Tryon, *op. cit.* See formulæ (11) and (15).

[4] R. C. Tryon, *op. cit.* See formulæ (13) and (14).

[5] The stipulation of equality of sigmas of the two fallible variables entering into th calculation of the reliability coefficient is clearly evident in the derivation of the Brown-Spearman formula. See C. Spearman, Correlation calculated from faulty data, *Brit. J. Psychol.*, 1910, 3, pp. 271–295. Note especially the proof in appendices *a* and *b* which leads to the important formula (9) on p. 290. Spearman here presents the

$\sigma_1 = \sigma_2 = \sigma_x$, say. Observe, now, the make-up of the fallible variances, by squaring, summing and averaging expressions (1) and (2):

$$\sigma_1^2 = \sigma_\infty^2 + \sigma_{e_1}^2, \tag{9}$$

$$\sigma_2^2 = \sigma_\infty^2 + \sigma_{e_2}^2. \tag{10}$$

Since $\sigma_1^2 = \sigma_2^2 = \sigma_x^2$, say, then $\sigma_{e_1}^2 = \sigma_{e_2}^2 = \sigma_e^2$, say. The variances of E_1 and E_2 being equal, then from (9), (10), (7), and (8)

$$\sigma_{e_1}^2/\sigma_1^2 = \sigma_{e_2}^2/\sigma_2^2 = r_{e_1x_1}^2 = r_{e_2x_2}^2 = \sigma_e^2/\sigma_x^2 = r_{ex}^2, \text{ say,} \tag{11}$$

$$\sigma_\infty^2/\sigma_1^2 = \sigma_\infty^2/\sigma_2^2 = r_{x_\infty x_1}^2 = r_{x_\infty x_2}^2 = \sigma_\infty^2/\sigma_x^2 = r_{x_\infty x}^2, \text{ say.} \tag{12}$$

That is, the per cent of X_1 produced by X_∞ equals the per cent of X_2 produced by X_∞, and also the per cent of X_1 produced by E_1 equals the per cent of X_2 by E_2.

What are the values of these two per cents, namely, of (11) and of (12)?

One more equation will lead us to these values. With a set-up as in Fig. 1, the following relation has been shown to be true [6]:

$$r_{12}^2 = r_{x_\infty x_1}^2 r_{x_\infty x_2}^2. \tag{13}$$

Then, from (12)

$$r_{12}^2 = r_{x_\infty x}^4,$$

$$r_{12} = r_{x_\infty x}^2 \tag{14a}$$

$$= \frac{\sigma_\infty^2}{\sigma_x^2}. \tag{14b}$$

The first answer to our question is in formula (14*a*, *b*) where r_{12}, *which is by definition the reliability coefficient, is the per cent by which* X *(either* X_1 *or* X_2*) is produced by* X_∞, this per cent expressible in terms of total correlation possible with X (14*a*) or in terms of total variance of X (14*b*).

famous split-test (odd-even elements) method of computing the reliability coefficient, a method which in the writer's opinion is most likely to insure that the discrepancies between fallible scores for the same individual are due to pure errors of measurement. See also, W. Brown, Some experimental results in the correlation of mental abilities, *Brit. J. Psychol.*, 1910, 3, pp. 296–322. In a footnote, p. 299, equality of sigmas is stipulated. See also T. L. Kelley, Statistical methods, Macmillan, 1923, p. 200.

[6] R. C. Tryon, *op. cit.* See formula (7) and also C. Spearman, The abilities of man, Macmillan, 1927, Appendix, p. iii.

From (3), (4), (11), and (12),

$$1 - r^2_{x_\infty x} = r^2_{ex}$$

and from (14*a*, *b*),

$$1 - r_{12} = r^2_{ex} \tag{15a}$$

$$= \frac{\sigma_e^2}{\sigma_x^2}. \tag{15b}$$

The second answer to our question is in formula (15*a*, *b*) where $1 - r_{12}$ *is the per cent by which* X *is produced by errors of measurement.*

Actual correlation between the fallible measure and its components appear immediately by taking the proper square roots. From (14*a*, *b*)

$$\sqrt{r_{12}} = r_{x_\infty x} \tag{16a}$$

$$= \sigma_\infty/\sigma_x. \tag{16b}$$

The square root of the reliability coefficient is the actual correlation coefficient of X with the true ability, X_∞, and it is also the ratio of the σ of the true ability to the fallible σ. As such, (16*b*) is Wright's path coefficient.[7]

Similarly, from (15*a*, *b*)

$$\sqrt{1 - r_{12}} = r_{ex}, \tag{17a}$$

$$= \sigma_e/\sigma_x. \tag{17b}$$

The square root of $1 - r_{12}$ is the correlation coefficient of the errors with the fallible score, and is the ratio [8] of σ_e to σ_x.

Besides contributing to the simplicity of understanding the reliability coefficient and its derivatives, the above interpretation of r_{12}, as the per cent of the fallible score produced by the true ability, indicates the proper basis for

[7] S. Wright, Correlation and causation, *J. Agricultural Research*, 1921, 20, p. 562.

[8] From (14*b*), the variance of X_∞ is $\sigma_x^2 r_{12}$, the square root of this quantity being the true sigma, σ_∞. From (15*b*), the variance of the errors is $\sigma_x^2(1 - r_{12})$, the square root of this being the well-known standard error of measurement formula. T. L. Kelley, Statistical methods, *op. cit.*, gives the derivation of $r_{x_\infty x}$, σ_∞, and σ_e in the chapter, 'Correlated measures,' but his approach is through the definition of X_∞ as being the average or sum of a score on an infinite number of tests like X_1. This makes the meaning of these formulæ nebulous, the derivations involved, and the notation confusingly complex. As derived above from the formulæ for percentage determination, they are simple in meaning and much briefer in derivation.

comparing and plotting different reliability coefficients. Where it is desired to plot the progressive change in reliability of a test through increase in the material in it (use of Brown-Spearman formula) or the change in reliability of the same test under differing experimental conditions, then the first power of the reliability coefficient should be plotted. The change in value of first power tells the change in per cent by which the true ability is operating in the fallible measurement, and no other function of r_{12} tells this. The use of $\sqrt{1 - r_{12}^2}$ to compare differences in reliability is misleading.[9] Employment of $\sqrt{1 - r_{12}}$ has some justification, since it is the coefficient of alienation when predicting from X_∞ to X, but such an expression is difficult to comprehend readily.

Application to the Correlation between Abilities

The above formulae enable one to ascertain definitely or within discoverable limits the percentage of determination of variable X by variable Y through the common factor, C, which occasions the correlation between them, as well as the percentage determination of each by their respective residual abilities (other than the common ability) and by their

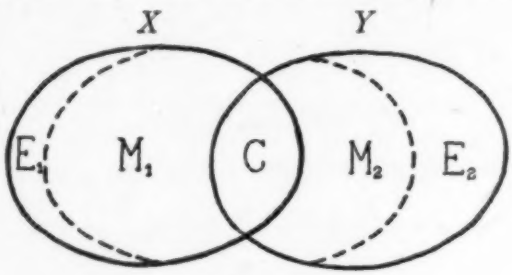

Fig. 2

respective errors of measurement. Diagrammatically, conceive the variables to be as in Fig. 2. C is defined as the *common ability*, the variance of which occasions the correlation between X and Y. M_1 is defined as the remaining *residual*

[9] In the papers by Ben D. Wood, Studies of achievement tests, *J. Educ. Psychol.*, 1926, 17, I: pp. 1–22; III: pp. 263–269, reliability coefficients were translated into $k = \sqrt{1 - r_{12}^2}$, and k's were compared for r's of test forms containing different numbers of items. Charts 1–4, and 12–22 contain plots of these k's and are rather misleading. The ordinates of these charts should have been on a scale of first power r's.

abilities in X independent of C. The variance of M_1 along with that of C make up the variance of X_∞. This is the same as saying that M_1 is the variance of X_∞ with the variance of C rendered constant. Similarly, M_2 is defined as the residual abilities in Y. E_1 and E_2 are the *errors of measurement* in X and Y respectively.[10] The subscripts, 1, 2, always refer to X and Y, respectively.

The following important questions are to be answered:

What is the percentage determination of fallible X and Y by their respective errors of measurement, E_1 and E_2, by the true abilities, X_∞ and Y_∞, by the common ability, C, and by the residual abilities, M_1 and M_2? What is the percentage determination of true X_∞ and Y_∞ by the common factor and by their respective residual abilities? What is the actual correlation between these variables?

All of these questions may be answered if one knows four expressions, the raw correlation between X and Y, r_{xy}, the reliability coefficient of X and Y, viz., r_x and r_y, respectively, and one determination, either r^2_{cx} or r^2_{cy}. The method of securing this last value will be discussed later.

The relationship between these various percentages must first be shown. The per cents by which X, for instance, is determined by its components, C, E_1, and M_1 must equal unity. That is,

$$r^2_{e_1x} + r^2_{m_1x} + r^2_{cx} = 1. \tag{18}$$

[10] The definitions amount to postulating that

$$x = h_1\left(\frac{c}{\sigma_c}\right) + k_1\left(\frac{m_1}{\sigma_{m_1}}\right) + l_1\left(\frac{e_1}{\sigma_{e_1}}\right), \tag{i}$$

$$y = h_2\left(\frac{c}{\sigma_c}\right) + k_2\left(\frac{m_2}{\sigma_{m_2}}\right) + l_2\left(\frac{e_2}{\sigma_{e_2}}\right). \tag{ii}$$

where the c's, m's, and e's are the variables as described above. The components of X being independent, then by squaring, summing and averaging, (i), and (ii) become

$$\sigma_x^2 = h_1^2 + k_1^2 + l_1^2, \tag{iii}$$

$$\sigma_y^2 = h_2^2 + k_2^2 + l_2^2. \tag{iv}$$

The squares of the h's, k's, and l's are the *effective variances* of the c's, m's and e's, respectively, in the composition of X. See R. C. Tryon, *op. cit.* In formulæ (11) to (15) of that paper, the percentage determinations are expressed in terms of effective variances. A pool of M_1 and E_1 were there lumped together and called D_x.

Similarly for Y,

$$r^2_{e_2y} + r^2_{m_2y} + r^2_{cy} = 1. \tag{19}$$

The proof of (18) has been relegated to the appendix.

The above questions are now answerable.

From (15*a*, *b*), as already proved,

$$r^2_{e_1x} = 1 - r_x, \tag{20}$$

$$r^2_{e_2x} = 1 - r_y. \tag{21}$$

The percentage determination of fallible X by its errors of measurements is unity less the reliability coefficient of X; and similarly of Y.

From (14*a*, *b*), as already proved,

$$r^2_{x_\infty x} = r_x, \tag{22}$$

$$r^2_{y_\infty y} = r_y. \tag{23}$$

The percentage determination of fallible X by its true scores is the value of the reliability coefficient of X; and similarly of Y.

The value r^2_{cx} presents the only real difficulty. Spearman has offered one method by which this value may often be secured. In certain other instances, the value is known on *a priori* grounds. When it is not secured by these methods, the *range* in which it must lie may be ascertained. We shall defer a discussion of this question until a later portion of this paper. For present purposes we shall assume that the value has been ascertained. Formula (13) and footnote 6 indicates that the following relation is true:

$$r^2_{xy} = r^2_{cx} r^2_{cy}. \tag{24}$$

Then

$$r^2_{cy} = \frac{r^2_{xy}}{r^2_{cx}}. \tag{25}$$

The percentage determination of X by C is r^2_{cx}, a value to be determined by Spearman's method or by a priori analysis. If not so available, the range within which it varies may be found. The determination of Y by C is r^2_{xy}/r^2_{cx}.

From (18) and (20),

$$\begin{aligned} r^2_{m_1x} &= 1 - r^2_{e_1x} - r^2_{cx} = 1 - (1 - r_x) - r^2_{cx} \\ &= r_x - r^2_{cx}. \end{aligned} \tag{26}$$

Similarly, and utilizing (25),

$$r^2_{m_2y} = r_y - \frac{r^2_{xy}}{r^2_{cx}}. \tag{27}$$

The percentage determination of X by its residual abilities is equal to the reliability coefficient less the per cent of X determined by C; and similarly of Y.

Since $x_\infty + e_1 = x$, by (1), then the correlation between C and X_∞ is that between C and X for a constant E_1, hence,

$$r^2_{cx_\infty} = r^2_{cx \cdot e_1} = \frac{(r_{cx} - r_{ce_1}r_{e_1x})^2}{(1 - r^2_{ce_1})(1 - r^2_{e_1x})}.$$

But $r_{ce_1} = 0$, hence $r_{ce_1}r_{e_1x} = 0$, and $(1 - r^2_{ce_1}) = 1$. Then employing (20),

$$r^2_{cx_\infty} = \frac{r^2_{cx}}{(1 - r^2_{e_1x})} = \left(\frac{1}{r_x}\right) r^2_{cx}. \tag{28}$$

Similarly, and utilizing (25),

$$r^2_{cy_\infty} = \left(\frac{1}{r_y}\right) r^2_{cy} = \left(\frac{1}{r_y}\right)\frac{r^2_{xy}}{r^2_{cx}}. \tag{29}$$

The percentage determination of X_∞ by the ability which it has in common with Y_∞, namely, C, is the reciprocal of the reliability coefficient of X times the percentage determination of X by C; and similarly of Y.

As with (3), the following relation holds:

$$r^2_{cx_\infty} + r^2_{m_1x_\infty} = 1. \tag{30}$$

Then, from (30), (28), and (26),

$$r^2_{m_1x_\infty} = 1 - r^2_{cx_\infty} = 1 - \left(\frac{1}{r_x}\right) r^2_{cx}$$

$$= \left(\frac{1}{r_x}\right)(r_x - r^2_{cx}) = \left(\frac{1}{r_x}\right) r^2_{m_1x}. \tag{31}$$

Similarly,

$$r^2_{m_1y_\infty} = \left(\frac{1}{r_y}\right)(r_y - r^2_{cy}) = \left(\frac{1}{r_y}\right)\left(r_y - \frac{r^2_{xy}}{r^2_{cx}}\right) = \left(\frac{1}{r_y}\right) r^2_{m_1y}. \tag{32}$$

The percentage determination of X_∞ by residual abilities other than the common factor is the reciprocal of the reliability coeffi-

cient of X times the percentage determination of X by M_1; and similarly by Y.

The correlation coefficients between the component variables and the fallible and true variables is secured by taking the proper square roots of the preceding formulæ (formulæ 20–32, excepting 24 and 30). The only correlation coefficient not apparent is that between X_∞ and Y_∞. This is the famous correction for attenuation formula presented by Spearman,[11] and proved by Yule.[12] It remains to bring this correlation into our determination system. Notice that the correlation between X_∞ and Y_∞ is simply that between X and Y with errors of measurement held constant. That is,

$$r_{x_\infty y_\infty} = r_{xy \cdot e_1 e_2} = \frac{r_{xy \cdot e_1} - r_{xe_2 \cdot e_1} r_{ye_2 \cdot e_1}}{\sqrt{1 - r^2_{xe_2 \cdot e_1}}\sqrt{1 - r^2_{ye_2 \cdot e_1}}}.$$

Since $r_{xe_2 \cdot e_1} = 0$, the product term in the numerator is zero, and the first root term in the denominator is unity, so that eventually, and from formulæ (20)–(23),

$$r_{x_\infty y_\infty} = \frac{r_{xy}}{\sqrt{1 - r^2_{xe_1}}\sqrt{1 - r^2_{xe_2}}} = \frac{r_{xy}}{r_{x_\infty x} r_{y_\infty y}} = \frac{r_{xy}}{\sqrt{r_x r_y}}. \qquad (33)$$

The determination, r^2_{cx}: The only term which the above formulæ require besides r_{xy} and r_x and r_y is the determination of X by C, to wit, r^2_{cx}. Where the common factor, C, runs through a number of variables besides X and Y, say through Z and W, and where the tetrad relations composed of the intercorrelations vanish, then r^2_{cx} may be secured by a formula presented by Spearman. This matter has been developed elsewhere and will not be discussed here.[13] Suffice it is to say that valid use of Spearman's formula requires that all the M's or residual abilities as well as errors of measurement be independent of each other.

[11] C. Spearman, The proof and measurement of association between two things, *Amer. J. Psychol.*, 1904, 15, p. 90.

[12] W. Brown, Some experimental results in correlation, *Comptes Rendus du VI^me Congrès International de Psychologie*, Genève, Aout, 1909. The proof was furnished by G. U. Yule.

[13] C. Spearman, Abilities of man, *op. cit.* See Appendix. Also, R. C. Tryon, *op. cit.* An illustration of the use of tetrad differences is given, as well as a short discussion of certain limitations in the use of them.

In many cases of correlation between X and Y an experimenter knows on *a priori* grounds that r^2_{cx} is a certain value. These special cases are given below, where it will be observed that in each case the above formulæ are greatly simplified. In all of the special cases the determination of X and Y by their respective E's and their respective true abilities remain unsimplified, *viz.*, $r^2_{x_\infty x}$ and $r^2_{y_\infty y}$ are always r_x and r_y respectively, and $r^2_{e_1 x}$, $r^2_{e_2 y}$ are always respectively, $1 - r_x$, $1 - r_y$. These determinations will not be mentioned again. Likewise, $r^2_{x_\infty y_\infty}$ remain unchanged.

(1) *The case in which X_∞ is the C factor itself, that is, where X_∞ is an element of Y_∞.* Such a structure is of frequent occurrence, *e.g.*, when X is an item in a composite test, Y, or where X is a given stage of learning in a total learning performance, Y. For this case,

$$r^2_{cx} = r^2_{x_\infty x} = r_x, \tag{34}$$

Hence, by reference to formulæ (25) to (33),

$$r^2_{cy} = \frac{r^2_{xy}}{r^2_{cx}} = \left(\frac{1}{r_x}\right) r^2_{xy} = r_y r^2_{x_\infty y_\infty}, \tag{35}$$

$$r^2_{m_1 x} = r_x - r^2_{cx} = r_x - r_x = 0, \tag{36}$$

$$r^2_{m_2 y} = r_y - r^2_{cy} = r_y(1 - r^2_{x_\infty y_\infty}), \tag{37}$$

$$r^2_{cx_\infty} = \left(\frac{1}{r_x}\right) r^2_{cx} = 1, \tag{38}$$

$$r^2_{cy_\infty} = \left(\frac{1}{r_y}\right) r^2_{cy} = r^2_{x_\infty y_\infty}, \tag{39}$$

$$r^2_{m_1 x_\infty} = \left(\frac{1}{r_x}\right)(r_x - r^2_{cx}) = 0, \tag{40}$$

$$r^2_{m_2 y_\infty} = \left(\frac{1}{r_y}\right)(r_y - r^2_{cy}) = 1 - r^2_{x_\infty y_\infty}. \tag{41}$$

Notice that in this special case the percentage determinations involve only the reliability coefficients and $r^2_{x_\infty y_\infty}$.

The opposite structure, in which Y_∞ is an element of X_∞, renders respectively formulæ (34) to (41) in which one reads X for Y.

(2) *The case in which C plays an equal role in X_∞ and Y_∞.* In Fig. 1, we observed such a structure as this. There X_∞ and Y_∞ *were C.* In his analysis of achievement and intelligence, Kelley hypothesized this special case to be the structure of all of the correlated variables in which he was interested.[14] For this case, $r^2_{cx_\infty} = r^2_{cy_\infty}$. But since $r^2_{cx_\infty} r^2_{cy_\infty} = r^2_{x_\infty y_\infty}$ [a formula of type (13)], then,

$$r^2_{cx_\infty} = r^2_{cy_\infty} = r_{x_\infty y_\infty}. \tag{42}$$

Hence by reference to formulæ (25) to (32), and (42),

$$\left(\frac{1}{r_x}\right) r^2_{cx} = r^2_{cx_\infty} = r_{x_\infty y_\infty}.$$

Then,

$$r^2_{cx} = r_x r_{x_\infty y_\infty}, \tag{43}$$

$$r^2_{cy} = r_y r_{x_\infty y_\infty}, \tag{44}$$

$$r^2_{m_1 x} = r_x - r_x r_{x_\infty y_\infty} = r_x(1 - r_{x_\infty y_\infty}), \tag{45}$$

$$r^2_{m_2 y} = r_y - r_y r_{x_\infty y_\infty} = r_y(1 - r_{x_\infty y_\infty}), \tag{46}$$

$$r^2_{m_1 x_\infty} = r^2_{m_2 y_\infty} = 1 - r_{x_\infty y_\infty}. \tag{47}$$

The percentage determinations in this special case [15] involve only the reliability coefficients and $r_{x_\infty y_\infty}$.

These above special cases of the general formulæ have derived from the various extreme and middle possibilities of the structure of X and Y. In many instances *a priori* analysis will lead one to the hypothesis that the relationship between a given X and Y represents one of these special cases. But for other instances in correlation where one has *no evidence* as to the composition of X and Y, and where Spearman's method of securing r^2_{cx} is not available, then the only valid procedure is to entertain the various extreme and middle possibilities of the composition of X and Y and thus determine the possible *range* of the production of X, Y, X_∞

[14] T. L. Kelley, Interpretation of educational measurements, World Book Co., 1927, Chap. 8, Sections 1 and 3.

[15] Another special case is that in which C plays an equal rôle in *fallible* X and Y, *i.e.*, $r^2_{cx} = r^2_{cy} = r_{xy}$ [by formula of type (13)]. The determinations come out to be: $r^2_{m_1 x} = r_x - r_{xy}$, $r^2_{m_2 y} = r_y - r_{xy}$, $r^2_{cx_\infty} = r_{xy}/r_x$, $r^2_{cy_\infty} = r_{xy}/r_y$, $r^2_{m_1 x_\infty} = 1 - (r_{xy}/r_x)$, $r^2_{m_2 y_\infty} = 1 - (r_{xy}/r_y)$.

and Y_∞ by the common cause and by their respective residual abilities.

An example: To illustrate the application of these methods to an actual problem, we shall use data presented by Kelley.[16] The problem is to discover the composition of the variables purporting to measure achievement and intelligence. The data were secured from a group of delinquent children who ranged over five school grades.

Let

X = *Achievement* (as measured by the Stanford Achievement Test),

Y = *Intelligence* (as measured by the Stanford-Binet).

The results as given by Kelley were as follows:

$$r_{xy} = .79, \qquad r_x = .98, \qquad r_y = .90, \qquad r_{x_\infty y_\infty} = .84.$$

We have here all of the information necessary for our determinations, except r^2_{cc}. On *a priori* grounds, there is no certain evidence that the structure of X and Y fits any one of the special cases above presented. The best procedure, therefore, is to draw up a table showing all of the special cases, thus examining the entire range of possible determinations. Such a table may be rapidly constructed as in Table 1. The determinations of X and Y by X_∞ and Y_∞, respectively, were not shown, because for all special cases and for the general case they are the same, namely, r_x [.98] and r_y [.90], respectively. The determinations of X and Y by their respective errors, E_1 and E_2, are the same for all cases, also, namely, $1 - r_x$ [.02], and $1 - r_y$ [.10], respectively.

The table is presented to show the simple interrelation between the various formulæ and to give an illustration of their use. Rows 5 and 7 are simply rows 1 and 3, respectively, with each term multiplied by r_x; similarly, rows 6 and 8 are respectively rows 2 and 4 with each term multiplied by r_y. Observe the *range* within each row. In row 1, for instance, the percentage determination of X_∞ by C, the common factor, ranges on one extreme (column 1) from unity or 100 per cent where X_∞ is C to the other extreme of $r^2_{x_\infty y_\infty}$

[16] T. L. Kelley, Interpretation of educational measurements, *op. cit.*, p. 207.

TABLE I

	General Formulæ	The special cases		
	Percentage determination of:	1. *Where all of X_∞ is in Y_∞*	2. *Where C equally determines X_∞ and Y_∞*	3. *Where all of Y_∞ is in X_∞*
1	*X_∞ by C* $r^2_{cx_\infty} = \left(\frac{1}{r_x}\right) r^2_{cx}$	1 [1.00]	$r_{x_\infty y_\infty}$ [.84]	$r^2_{x_\infty y_\infty}$ [.71]
2	*Y_∞ by C* $r^2_{cy_\infty} = \left(\frac{1}{r_y}\right) \frac{r^2_{xy}}{r_{cx}}$	$r^2_{x_\infty y_\infty}$ [.71]	$r_{x_\infty y_\infty}$ [.84]	1 [1.00]
3	*X_∞ by M_1* $r^2_{m_1 x_\infty} = \left(\frac{1}{r_x}\right)(r_x - r^2_{cx})$	0 [.00]	$1 - r_{x_\infty y_\infty}$ [.16]	$1 - r^2_{x_\infty y_\infty}$ [.29]
4	*Y_∞ by M_2* $r^2_{m_2 y_\infty} = \left(\frac{1}{r_y}\right)\left(r_y - \frac{r^2_{xy}}{r^2_{cx}}\right)$	$1 - r^2_{x_\infty y_\infty}$ [.29]	$1 - r_{x_\infty y_\infty}$ [.16]	0 [.00]
5	*X by C* r^2_{cx} = (see text)	r_x [.98]	$r_x r_{x_\infty y_\infty}$ [.82]	$r_x r^2_{x_\infty y_\infty}$ [.70]
6	*Y by C* $r^2_{cy} = \frac{r^2_{xy}}{r^2_{cx}}$	$r_y r^2_{x_\infty y_\infty}$ [.64]	$r_y r_{x_\infty y_\infty}$ [.76]	r_y [.90]
7	*X by M_1* $r^2_{m_1 x} = r_x - r^2_{cx}$	0 [.00]	$r_x(1 - r_{x_\infty y_\infty})$ [.16]	$r_x(1 - r^2_{x_\infty y_\infty})$ [.28]
8	*Y by M_2* $r^2_{m_2 y} = r_y - \frac{r^2_{xy}}{r_{cx}}$	$r_y(1 - r^2_{x_\infty y_\infty})$ [.26]	$r_y(1 - r_{x_\infty y_\infty})$ [.14]	0 [.00]

(column 3) where Y_∞ is C. These are the upper and lower *limits* by which X_∞ may be determined by C. For the middle case (column 2) where C equally determines X_∞ and Y_∞, the per cent is $r_{x_\infty y_\infty}$. Correspondingly, the per cent by which Y_∞ is determined by C in *row* 2 ranges from the lowest possible per cent, $r^2_{x_\infty y_\infty}$, through $r_{x_\infty y_\infty}$ to unity or 100 per cent. In our example, the factors that are identical (C) in achievement (X) and intelligence (Y) range from a possible 71 per cent ($r^2_{x_\infty y_\infty} = .71$ in brackets in the table) to 100 per cent. If identical factors play an equal role in intelligence and achievement they produce 84 per cent ($r_{x_\infty y_\infty} = .84$). In this connection, it must be pointed out that Kelley has considered only this middle possibility in his book 'The Interpretation of Educational Measurements' where he has endeavored to show that there is a high community of function between intelligence and achievement (chap. 8). The square of $r_{x_\infty y_\infty}$ represents, however, the least possible community of function, and not the first power.[17]

Rows 3 and 4 give the possible range of determination of X_∞ and Y_∞ by residual abilities, and it may be seen that this range is between zero and $1 - r^2_{x_\infty y_\infty}$. In our example, this means that true intelligence and true achievement may be determined any degree from 0 to 29 per cent ($1 - r^2_{x_\infty y_\infty} = .29$) by residual abilities. For the greatest determination, residual abilities might correlate as high as $\sqrt{.29}$ or .54.

The determination of fallible intelligence and fallible achievement by the common factor is given in rows 5 and 6. The range is between the value of the reliability coefficient and the product of the coefficient by $r^2_{x_\infty y_\infty}$. For X this lies between 98 per cent and 70 per cent, for Y between 90 per cent and 64 per cent. In rows 7 and 8 are the determinations of the fallible scores by the residual abilities. For our example, the determination of X by M_1 may be between 0 and 28 per cent, of Y by M_2 between 0 and 26 per cent.

[17] This is not intended as a criticism of Dr. Kelley's analysis. His assumption that C plays an equal rôle in X_∞ and Y_∞ is one which he admits may be in error (p. 195). The writer desires only to point out the *range* of percentage possibilities, from which Dr. Kelley has selected only the middle case. He gives only the one determination of the $r_{cx_\infty}{}^2$ type and interprets it only in terms of variances.

It is important to remember that one really is dealing in any experimental situation with *fallible* X (or Y) scores, and that the percentage determination of these fallible X scores by C is r_{cx}^2 and by residual abilities *plus* errors of measurement is $1 - r_{cx}^2$. In our example, the fallible actually obtained intelligence scores may at the lowest extreme be determined only 64 per cent by achievement (through C). This means that 36 per cent may come from other factors, which as a pool could correlate to the extent of $\sqrt{.36}$ or .60 with Y. Regarding the middle case, the fallible intelligence scores have only 76 per cent in common with achievement, leaving for a pool of other factors than C to produce 24 per cent of fallible intelligence and to correlate with Y to the extent of $\sqrt{.24}$ or .49. The true intelligence measures (*i.e.*, fallible intelligence with errors of measurement removed) as Kelley showed, were produced 84 per cent by the common factor, leaving residual abilities the possibility of producing 16 per cent of true intelligence and correlating $\sqrt{.16}$ or .40 with it.

Assumptions: The assumptions underlying the use of these methods are the same as those mentioned in the former paper.[18] *No assumptions* have been made as to the *nature* of the elements which produce the errors of measurement, residual abilities, or common factor. All of the derivations have been based upon the *variance* of the abilities involved and of the *degree of correlation* between the abilities. In practice, care must always be taken that the reliability coefficients used satisfy the definition of the reliability coefficient, and the correlations between the variables must, of course, be linear, or by proper transmutation of scores made to be linear. One basic assumption has been made, the *additive hypothesis*, by which it is supposed that the component variables produce the fallible and true variables by summation. As Spearman has pointed out,[19] this hypothesis was used by Bravais in deriving the original expression involving the correlation coefficient. Further, even though the composite variable *is* a function of more complicated inter-action of components than summation, according

[18] R. C. Tryon, *op. cit.* See Section III.

[19] C. Spearman, Abilities of Man, *op. cit.* Appendix xv.

to Spearman, the function may be mathematically reducible by Taylor's theorem to approximately the summation we have hypothesized. Where these demands are not satisfied, the percentage determinations and correlations secured by these methods must only be considered to be approximations to the true state of affairs. No new assumptions have been added in the derivation of these formulæ than those which lie at the basis of the derivation of the reliability coefficient and of the correction for attenuation formula.[20]

Appendix

Proof that: $r^2_{e_1x} + r^2_{m_1x} + r^2_{cx} = 1$.

Give the variables numerical notation, as follows: $X(4)$, $E_1(1)$, $M_1(2)$, and $C(3)$. It is given by postulation that X is completely determined by E_1, M_1, and C, and that all the intercorrelations between these three components equal zero.

Hence the multiple correlation coefficient:

$$R_{4(1,\,2,\,3)} = \sqrt{1 - k^2_{14}k^2_{24.1}k^2_{34.12}} = 1. \qquad \text{(i)}$$

Squaring, and expanding $k^2_{24.1}k^2_{34.12}$,

$$R^2_{4(1,\,2,\,3)} = 1 - k^2_{14}(1 - r^2_{24.1} - r^2_{34.12} + r^2_{24.1}r^2_{34.12}) = 1. \qquad \text{(ii)}$$

By writing partial coefficients, and properly manipulating r^2's which equal zero, then

$$r^2_{24.1} = \frac{r^2_{24}}{k^2_{14}} = B, \quad \text{say,}$$

$$r^2_{34.12} = \frac{r^2_{34}}{k^2_{14}(1 - B)}.$$

Substituting in (ii),

$$1 - k^2_{14}\left[1 - B - \frac{r^2_{34}}{k^2_{14}(1 - B)} + B\frac{r^2_{34}}{k^2_{14}(1 - B)}\right] = 1,$$

$$1 - k^2_{14}\left[(1 - B)\left(1 - \frac{r^2_{34}}{k^2_{14}(1 - B)}\right)\right] = 1,$$

$$1 - k^2_{14}\left[1 - B - \frac{r^2_{34}}{k^2_{14}}\right] = 1.$$

[20] In a recent book, Crossroads in the mind of man, Stanford University Press, 1928, pp. 1–238, Dr. Kelley gives an admirable chapter, 'Theory and technique,' which shows the correlational conditions which must obtain if a number of variables are to be analyzed into one or more group factors plus residuals. On page 36, I find that he mentions our formulæ (14*b*) and (15*b*). His formula (5) is also our formula (37). He likewise gives a percentage interpretation to these particular determinations, but expresses them only in terms of variances.

Substituting the value for B, and multiplying through the bracket by k_{14}^2, one gets eventually,

$$r_{14}^2 + r_{24}^2 + r_{34}^2 = 1. \qquad \text{(iii)}$$

Summary

The percentage degree of determination of one variable by another which is correlated with it is defined either as the per cent of variance (σ^2) of one variable contributed by the variance of the other or as the per cent of total possible correlation of one variable (namely unity) contributed in terms of correlation by the other. Utilizing these definitions, the numerical value of the reliability coefficient, r_x, of a variable, X, is the per cent of fallible X produced by the true ability, X_∞. The per cent contributed by errors of measurement is $1 - r_x$.

When two variables, X and Y, are correlated, by utilizing only the values, r_x, r_y, r_{xy}, and one other term, r_{cx}^2, where C is the variable occasioning the correlation between X and Y, one may ascertain the percentage determination of each fallible variable by its component errors of measurement, true ability, common ability, and residual abilities apart from the common ability. The per cents by which each true ability is determined by the common ability and by its residual abilities may likewise be found. The value, r_{cx}^2, may be computed in some cases by use of Spearman's tetrad difference method, or in other cases it may be known on *a priori* experimental grounds. In instances where the value of it is not obtainable, then by utilizing the *range* within which it may fall, one may secure the range within which each of the percentage determinations may fall. The formulæ for the highest and lowest possible per cents as well as for the middle per cent are for any of these determinations some simple function of the reliability coefficient or of the correction for attenuation formula, $r_{x_\infty y_\infty}$, or of both of these expressions.

[MS. received August 6, 1929]

INTELLIGENCE TESTS OF IMMIGRANT GROUPS

BY CARL C. BRIGHAM

Princeton University

A number of articles and books have appeared reporting the results of the psychological examination of immigrants in the United States. In the present article, the general trend of the current studies will be summarized or interpreted without extensive citation of references. This summary may be regarded as in the editorial form, and is not a systematic review of the literature on the subject. The articles in question are practically all in the following form: the Blank test of intelligence was given to n children in x grades of a certain school system; the average test scores of children of nationality A, B, and C were as follows. Our inquiry will disregard questions of sampling, methods of treating data, and the like, and will direct itself to the more fundamental problem of the meaning of test scores.

Psychologists have been attacked because of their use of the term 'intelligence,' and have been forced to retreat to the more restricted notion of *test score*. Their definition of intelligence must now be 'score in a test which we consider to measure intelligence.' The ordinary test includes several different types of sub-tests for which varying time allowances are made. Each test battery, or collection of sub-tests, is a closed system which may be studied in its own right, first by the method of correlation with independent criteria, and second by various internal criteria of reliability and self-consistency.

The diversion of psychologists from general debates on the question of the nature of intelligence to the more careful analysis of test scores has resulted in several important advances. The question immediately arises as to whether or not the test score itself represents a single unitary thing. This question is crucial, for if a test score is not a unitary thing it should not be given one name.

Professor Spearman of the University of London has for many years advocated the two factor theory of intelligence, which is to the effect that there is a general factor, '*g*,' underlying performance in sub-tests of diverse sorts, and special factors, '*s*'s,' unique to each type of sub-test. Two similar sub-tests within one test battery may give rise to a 'group factor' which is not *g*. Earlier discussions concerning the validity of this hypothesis revolved around a technical issue on the validity of a certain method of proof. Another method has more recently been advanced by Spearman, and this method has been elaborated by Professor Kelley of Stanford University. The present discussion in this field is no longer concerned so much with the validity of Spearman's theory as it is with the proper statement of the number of factors into which a test score may be resolved. The issues here are very involved, as it has been found that the correlations between sub-tests depend on time allowances, and on the difficulty and validity of single test items making up the sub-tests. Until some of the variables are more adequately controlled, no satisfactory answer will be found.

While these investigations are in progress and no guess can yet be made as to their probable outcome, there are many findings which effectively challenge certain interpretations of test results made in the past. For the sake of illustration, one may consider that an 'intelligence' test battery containing six sub-tests and an 'achievement' test battery also containing six sub-tests have been given to the same group of people. If one should regard this as one battery of tests, and should study all 66 of the possible bonds between pairs of sub-tests, one could hardly expect to find the dichotomy in the names of the two tests revealed in such an analysis of their nature. A safe procedure is to disregard what a psychologist calls his test, and to study the test itself as an instrument more or less adequate to its purpose. Most psychologists working in the test field have been guilty of a *naming fallacy* which easily enables them to slide mysteriously from the score in the test to the hypothetical faculty suggested by the name given to the test. Thus, they speak of sensory

discrimination, perception, memory, intelligence, and the like while the reference is to a certain objective test situation.

A far-reaching result of the recent investigations has been the discovery that test scores may not represent unitary things. It seems apparent that psychologists in adding scores in the sub-tests in some test batteries have been doing something akin to adding apples and oranges. A case in point is the army alpha test.

All of the methods of analyzing the army test data were evolved from a statistical study of a special experimental group of 1047 cases. Spearman [1] has analyzed the army test data and finds that they do not fit his hypothesis. His explanation follows:

"In the endeavour to procure fairly representative samples of the Army, the plan had been conceived of sending round to twelve different camps, asking each that the tests should be given to 'approximately 100 relatively unselected men.' Nine of the camps responded, their contributions of men ranging from under 20 to over 200. But by such a procedure the subjects, and still more the testing, must have become heterogeneous to the last degree. For instance, when in any camp the testing or marking happened to be more generous than in others with respect to any of these tests, then the men here would tend to shine in these particular tests; the result must be to generate additional correlation between these tests quite independently of any resemblance or psychological connection between them. And the cumulative result of such influences could scarcely fail to produce the highly irregular character that we actually find" (p. 157).

Kelley [2] has criticized the above interpretation as follows:

"The data which most adequately take into account maturity and sex is the Army Alpha data quoted very disparagingly by Spearman. As to the facts we can all agree that a single general factor is not indicated. As to the cause, Spearman writes that by the procedure followed, 'the subjects, and still more so the testing, must have become heterogeneous

[1] C. Spearman, The abilities of man, New York, Macmillan, 1927, pp. 415 + xxxiii.

[2] T. L. Kelley, Crossroads in the mind of man, Stanford University, Calif., Stanford University Press, 1928, pp. 238.

to the last degree.' Knowing something of the care with which this testing work was done, the writer does not believe the second of these charges is justified. As to the first charge, it is true that the group was rather heterogeneous, but this fact would introduce a *g* factor, not take out one already there. In brief, these data, almost unique in that they allow for maturity and sex, for the group consisted of adult men, do not yield a comprehensive *g* factor " (p. 19).

There were eight tests in the army alpha battery as follows:

1. Directions
2. Arithmetical reasoning
3. Practical judgment
4. Synonym-antonym
5. Disarranged sentences
6. Number series
7. Analogies
8. Information

There are certain difficulties involved in the analysis of the army test data which have been summarized adequately by Kelley:

"The special study of the inter-correlations between the eight sub-tests of the Army Alpha Test, conducted by the Division of Psychology of the Surgeon General's Office (Yerkes, 1921), provides data upon a large population of American white English-speaking adults in nine different army camps. The special merits of these data are that they cover a large population, sex is constant, and maturity substantially constant. The defects are that racial homogeneity is not attained, the tests separately have low reliability, and a technique of treatment was followed which makes it impossible to calculate the probable errors of the correlation coefficients or of the tetrad differences. The number of cases examined was 1047, and it was found that the distribution of a number of sub-tests gave undistributed extreme scores, zero, or perfect scores. Accordingly, all correlations were corrected for these undistributed scores, to obtain estimates of the correlation

maintaining if the tests had permitted lower or higher scoring as required. Though the writer sees no reason to believe that these very carefully made corrections either introduced or took out any general or group factors, still it is true that the corrections are of such magnitude that ordinary probable-error formulas may not be used for the determination of reliability" (pp. 215–216).

Kelley has analyzed the army test data from the special experimental sample and has designated in words the probable nature of bonds found between pairs of tests. He finds what he calls a numerical factor common to tests 1 and 2, tests 1 and 6, and tests 2 and 6; a verbal factor common to tests 3 and 4, tests 3 and 8, and tests 4 and 8. He also notes a factor common to tests 5 and 7, and a factor common to tests 6 and 7.

The present writer has made a similar analysis introducing the Stanford-Binet examination as another test (called 9) in the series. This study shows the same bonds noted above with the addition of another bond (5–8) which Kelley might have included, and another possible bond (4–9) about the size of the 1–2 bond. Because of certain inadequacies in such studies arising from the defects of the original data, and the methods which it was necessary to use in calculating the coefficients of correlation between the sub-tests, it is, perhaps, not worth while to attempt a more systematic study of the data from the army alpha test. After puzzling over the situation longer than the data warrant, the writer offers the following not entirely satisfactory solution.

The assumption is made that people taking the alpha test adopted two different attitudes or sets, viz., a 'directions attitude,'—an attitude of careful attention to the examiner's instructions without looking at the test questions while the directions were read; and a 'reading attitude,'—partially or completely ignoring the examiner's instructions while studying the test questions during the time in which the examiner was reading. The adoption of the first attitude would tend to give the individual higher scores in test 1 (entirely oral directions), test 6 (an unusual form of mathematical test),

and in test 7 (a novel type of verbal test). On the other hand, a person adopting the second attitude might quickly find out what was required in tests 3, 4, 5, and 8, and his score would be better if he ignored what the examiner was reading and studied the test questions during the period of instruction. The adoption of this reading attitude, while tending to make scores higher in tests 3, 4, 5, and 8, might make for lower scores in tests 1, 6, and 7. The writer's analysis shows nothing inconsistent resulting from the combination of tests 1, 6, 7, and 9 (the Stanford Binet), nor from the combination of tests 3, 4, 5, and 8, while all of the sixteen possible bonds between the two groups of four tests are negative except 4–9 and 5–7. A fairly good case may therefore be made for the directions attitude in tests 1, 6, 7, and 9, and for the reading attitude in tests 3, 4, 5, and 8. There is also evidence of what might be guessed to be a mathematical factor in tests 1, 2, and 6, the directions attitude making for still higher scores in tests 1 and 6. The mathematical factor in tests 1, 2, and 6 is independent of the 3, 4, 5, and 8 complex, and all twelve cross bonds are negative here except 2–3 which is slightly positive. The only remaining positive bonds not accounted for appear between tests 5 and 7, and between tests 4 and 9.

The report of the army examiners[3] shows in Table 158 (p. 640) that scores in different sections of the alpha scale were derived from different tests. One may estimate from this table, for example, that tests 4 and 7 would rarely enter into a total score around 30 points, while persons scoring at the upper end of the scale, as high as 190, would normally get 60 or 70 points from these two tests. At the various levels of scores, different sub-tests contribute to the total in varying amounts so that alpha was not a consistent test even from this standpoint. This fact together with the facts previously noted which suggest disparate group factors within the test show that the eight tests of army alpha should not have been added to obtain a total score, or, if added, similar total scores should not have been taken to represent similar performances in the test.

[3] Memoirs of the National Academy of Sciences, Vol. XV.

If the army alpha test has thus been shown to be internally inconsistent to such a degree, then it is absurd to go beyond this point and combine alpha, beta, the Stanford-Binet and the individual performance tests in the so-called 'combined scale,' or to regard a combined scale score derived from one test or complex of tests as equivalent to that derived from another test or another complex of tests. As this method was used by the writer[4] in his earlier analysis of the army tests as applied to samples of foreign born in the draft, that study with its entire hypothetical superstructure of racial differences collapses completely.

Kelley,[5] in compiling a list of traits which he has found to be more or less independent of each other, enumerates:

"(1) Verbal intelligence, or the ability which in the main underlies facility in naming opposites, coördinates, subordinates, supraordinates, predicates; and found in tests of mixed relationships, practical judgments, vocabulary, written directions, sentence completion (textual matter of literary content), sentence meaning, paragraph meaning, word meaning, and logical selection.

"(2) Quantitative intelligence, or the ability in the main underlying facility in computation and other situations involving numbers as content.

"(3) Spatial intelligence, or the ability in the main underlying facility in handling form boards, geometrical forms, and right- and left-hand, Knox cube, and other similar tests" (p. 122). He states that verbal intelligence "can be determined by school achievement tests of the reading and vocabulary sorts and by general intelligence tests or such portions of them as are of a verbal nature and do not include numbers and spatial relations as content" (p. 124).

These findings of Kelley as to the probable existence of independent traits rather effectively challenge current test procedure, as most tests combine verbal and mathematical materials indiscriminately. As most of the non-verbal tests

[4] C. C. Brigham, A study of American intelligence, Princeton, N. J., Princeton University Press, 1923, pp. 210.

[5] T. L. Kelley, Interpretation of educational measurement, Yonkers, N. Y., World Book Co., 1927, pp. 363.

deal with spatial relations, one should not attempt to find similarities between scores in verbal and non-verbal tests. The writer does not hold that studies of the independence of traits have advanced far enough to have final theoretical significance. It is, however, apparent that scores in tests found to be independent should not be added.

For purposes of comparing individuals or groups, it is apparent that tests in the vernacular must be used only with individuals having equal opportunities to acquire the vernacular of the test. This requirement precludes the use of such tests in making comparative studies of individuals brought up in homes in which the vernacular of the test is not used, or in which two vernaculars are used. The last condition is frequently violated here in studies of children born in this country whose parents speak another tongue. It is important, as the effects of bilingualism are not entirely known.

This review has summarized some of the more recent test findings which show that comparative studies of various national and racial groups may not be made with existing tests, and which show, in particular, that one of the most pretentious of these comparative racial studies—the writer's own—was without foundation.

[MS. received September 7, 1929]

SCIENCE AND CULT[1]

BY W. S. TAYLOR

Smith College

If from the vantage-ground of the armchair a psychologist were suddenly required by a searching deity to distinguish between the sheep and the goats in the scientific field, by what signs would he know them? It is indeed a commonplace that that heavenly dichotomy would be too simple for our more mundane sphere. Pure sheep are very hard to find, and pure goats are almost as rare. Many individuals have characteristics of both sheep and goats. Some very good sheep are goats upon occasion. Some apparent sheep have goatish souls, just as some that appear to be goats are, spiritually, sheep. Furthermore, goats sometimes lose their Egos and find Reality. Leaving aside however this problem of what our judge would have to do with all the scientific animals that come between, this paper will attempt idealistic norms for the extreme types of these intellectual companies.

Science and cult differ fundamentally in their intellectual attitude. The intellectual attitude of the pure scientist is, characteristically, hypothesis; the attitude of the pure cultist is unhypothetical conviction. In other words, the scientist entertains suppositions as 'candidates for verification,' as Dotterer puts it; while the cultist holds to propositions as true already. Yet it is worthy of note that the scientist is not paralyzed by doubt, any more than the cultist is—if as much, since the cultist is often almost paralyzed by the danger of doubt; while waiting for his newer suppositions to be chosen or rejected by truth, the scientist assumes as workable those better established hypotheses which are called facts. Thus the scientist, though living always in the presence of possibility, lives and works on the hearty basis of probability.

[1] Read at the Ninth International Congress of Psychology held at New Haven, Conn., September 4, 1929.

It follows, however, that whereas the scientist is confident that his ideas may be changed any time by facts, the cultist is confident that his ideas, being inherently true, are in no danger of change.

The scientist, moreover, has some perspective with regard to allied fields of investigation: he knows that there are allied fields, fields as important as his own; and he knows that allied or adjacent fields naturally support one another. The cultist, on the other hand, must be blind to his natural allies, claiming at least that he needs no support from them, or that they can provide no support for his intellectual completeness. Similarly, with respect to history, the scientist is aware that his field of study is a growth, with a heritage from many different workers; while the cultist usually thinks that his creed had a beginning, that it sprang full-blown into its proper shape. If the real cultist acknowledges any measure of growth in his doctrine, it is at most only a progressive growth, never a progressive revision.

The scientist's devotion, ideally, is to truth; while the cultist's interest is to establish his own position. Thus the scientist's hypothesis is outside himself, so to speak; and when you present data sufficient to change that hypothesis, the pure scientist welcomes the change as an interesting inconvenience. The cultist's doctrine, in contrast, is as it were within himself; so that when you doubt the truth of his view, the cultist is hurt, unless he ascribes your doubt to sin or ignorance or both. In other words, instead of being interested and open-minded with regard to criticism, the cultist becomes self-defensory. If the cultist enters into intellectual argument at all, then like all controversialists he tends to exaggerate the difference between his opponents' position and his own, absurdly distorting or over-simplifying the views he attacks and fears. No wonder it is the scientific spirit that is naturally international, as well as truly liberal.

It follows that with respect to his pupils, the scientist is a teacher; the cultist is a preacher.

What conditions are most favorable to cults?

Obviously the simplicity, the inertia, of human thinking

is a factor. We all tend to believe dogmatically whenever we can; we hypothesize when we must. This simplicity is related again to our tendency to follow leaders and crowds, to be protected by an intellectual parent or gang. Further, a prime condition that motivates the cult is the presence of human values, great needs or desires. (This is why a cult is usually organized about some principle which has potential scientific value.) A final condition most favorable to the life of the cult is the lack of experimental check, controlled observation.

A little survey of those areas in which cults are most noticeable seems to substantiate these conditions. The provinces of mechanics, physiology, general medicine, psychotherapy, politics, religion, represent a scale both of increasing values and of decreasing experimental check; so that from the relative cultlessness of mechanics to the extreme cultishness of the province of religion, we have an increasing appeal to the human tendency to accept without proof. Psychotherapy, in particular, as applied abnormal psychology, seeks to control conditions of the enjoyment of life itself, yet greatly lacks experimental precision. In these ways psychotherapy, and with it abnormal psychology, are predisposed to cultishness. From cultishness, however, we should like to see every field of knowledge free.

The remedy would seem to lie in more facts; in more of the attitude of hypothesis; and in more checks on our interpretations—experimental checks especially, but when experimental checks are lacking, then we need logical checks. These are developed best in free exchange of ideas.

[MS. received October 1, 1929]

MOTIVE AS A CONCEPT IN NATURAL SCIENCE *

BY FLOYD H. ALLPORT

Syracuse University

The concept of motivation suffers from a vagueness which has long been a stumbling block to psychological research. This paper will attempt to show the reasons for this situation and to indicate the methodological possibilities and limitations in the study of motivated behavior.

For this purpose all objects proposed for scientific study may be divided into two groups distinguished by the type of behavior of the scientist by which they are identified. We have, respectively, objects capable of *explicit denotation*, and objects (perhaps it would be better to call them entities) which can be made known only through *implicit denotation*. In explicit denotation the object is pointed out by some form of skeletal response. We are also able to react successively, either by pointing, touching, or manipulation, to a series of points lying within it. Implicit denotation, on the other hand, is not a point-for-point response of the scientist's organism to the stimulus field, but consists of descriptive or symbolic reactions made about the object in verbal or other terms. A wire, for example, exhibiting certain changes as an electric current passes, is *explicitly* denotable; the units in which those changes are recorded, as well as the so-called energy of the current itself, are capable only of *implicit* denotation. A piece of iron is explicitly denotable; its alleged atoms and electrons are restricted, as yet, to implicit denotation. A nerve fiber as it conducts an impulse is capable of explicit denotation; conduction conceived as a process is identifiable only implicitly. A neuron chain is explicit, a habit implicit, and so on. Now with regard to motivated behavior—it is my thesis that it is only the organism (or its

* Read at the Ninth International Congress of Psychology, held at New Haven, Conn., September, 1929.

parts) as it moves, or changes position, which can be treated by the explicit procedure.[1] Drive, instinct, incentive, goal-value of the object, and motive are, like atoms, kilowatts, and processes, capable only of implicit denotation.[2]

To reveal the point of this contrast let us compare the situation in motivated behavior with that in gravitation. According to the Newtonian observation (and that is sufficiently exact for our purposes) two objects tend to move toward each other at a rate which varies directly with the product of their masses and inversely with the square of their distances apart. If we hold the mass of two bodies constant but decrease their distance from one another we increase accordingly the degree of their tendency to move toward one another.[3] We have here, as in all scientific laws, simply a statement of the relation between two variables, each determinable by explicit denotation of the objects concerned. So far we have projected no cause or force into the situation. The word gravitation, if used at all, is used only as a convenient synonym for the descriptive mathematical statement.

Now let us imagine a rat which has not been fed for twenty-four hours brought into a situation in which he is shut off from access to food only by a problem, such as a concealed doorway or simple multiple choice, which is for him soluble. We note that his movements in this situation are relatively rapid and frequent, and that consequently he may solve the problem more quickly than if he had recently been fed. Here again we have a constant relation between two variables, the contractions of muscles in the rat's stomach and the increase in rate of skeletal movement, two variations which can be determined, directly or indirectly, through

[1] The stimulus object, provided it is not thought of as a 'meaningful situation' is, of course, also explicitly denotable.

[2] These terms should not be confused with 'implicit and explicit *behavior*' as used by certain behaviorists. The latter expressions refer only to the organism itself, with regard to whether the behavior is hidden from view or manifest. Explicit and implicit *denotation* have a very different meaning in that they indicate certain relationships between the scientist's behavior and the objects to which he reacts (*i.e.* which he studies).

[3] The fact that, owing to the great mass of the earth always present in the situation, this demonstration is possible only upon bodies of a planetary order does not invalidate the principle involved.

explicit denotation of the objects concerned. Holding other conditions constant it might even be possible to secure an approximate mathematical statement of this relationship. Here also, as in the case of gravitation, we need assume no force or drive. Words like hunger, instinct, and motive, if used at all, are merely less exact names which we use for the co-variation above described. We do not think of motive as a force in the rat's organism driving him to make rapid movements which eventually bring him to the food, any more than we posit in the earth or the stone a force which propels them toward each other.

This formulation, however unsatisfactory it may be to psychologists of the hormic school, has one clear advantage. It makes possible what I shall call a *multi-level approach* to the phenomena we are studying. That is to say, it enables us experimentally to break up the objects studied into their components, and to discover what goes on at this simpler level concomitantly with the events which transpire at the more complex. The advantage of this procedure, in turn, is that it enables us to reach more basic generalizations, to explain the exceptions which have appeared in our previous formulations, and to predict not only the familiar but the unexpected happening. To take an example, we may note that certain exceptions have appeared even in a law as apparently stable as that of gravitation. A possible approach to this problem is, by experiment, to analyze our materials (at least conceptually) into smaller particles and see what happens at the more elementary levels. By this method modern physicists have, in fact, begun to catch a glimpse of more basic laws which comprise under one formulation the old law of gravitation and certain phenomena, such as those of electromagnetism, which were previously regarded as unaccountable exceptions. Knowing such a formulation, we might be able actually to predict events which would otherwise be quite unintelligible.

Similarly, in the case of the rat, should a sudden anomaly occur in the curve of motivated learning, such anomaly can be made the strategic point for analysis to determine what

is happening at the neurological and the physico-chemical levels of the animal's structure. There might, for example, be discovered products of catabolism or hormone secretions which would throw light both upon the disturbance and the usual process, and would enable us to combine both the familiar law and its exception in a more fundamental generalization.

This attacking of exceptions through the multi-level approach has been of great importance throughout the history of scientific discovery. And it is my thesis that the multi-level approach, with its peculiar advantages, is made possible only by keeping before us as our primary objects of study things which are capable of *explicit denotation.* In cases where only implicit identification is possible these benefits are lost.

But even though the view of motive as an implicitly denoted force must forego these advantages, is it therefore useless? By no means. The engineer finds it helpful to think of gravitation and electrical energy as definite working forces. Similarly, the applied psychologist can use categories of instincts, desires, and motives in controlling men for practical rather than scientific ends. Exceptions to these formulations, however, cannot be dealt with. The telic, implicit approach, yields only uni-level observations. If the hunger-motive fails to operate in five trials out of a hundred, the telicly conceived law of hunger as a drive of behavior must forever remain only ninety-five per cent perfect. No wider generalization can be secured until the objects are re-approached with the attitude of explicit, rather than implicit, denotation.

But in spite of the advantages described for the explicit method, there must be pointed out, for this procedure also, some genuine limitations. It can give us laws of learning in general, but that the rat will learn this particular combination of problem-box habits as a part of his neural organization cannot be predicted in any scientific fashion. The physicist can state laws for all falling bodies alike, but he cannot accurately foretell what bodies will fall, or when and how far,

or what the pattern of their combined falling will be. Again, the qualitative aspect of motive must, in the explicit approach, be completely lost. In industrial workers, for example, diverse situations, such as those of resentment, ambition, or family affection, may have afforded conditioning stimuli for the same type of visceral response. The method of explicit denotation, which deals only with bodily changes at the present instant, will overlook these differences of motive which, from the subjective standpoint, are indubitable. The genetic as well as the telic perspective is lost when we cut across the life-stream by the approach of explicit, multi-level analysis.

To insist, therefore, that all aspects of motivation can be illuminated through explicit denotation is to close our eyes to the limits of our method. To expect, on the other hand, that an implicit, telic, or anthropomorphic denotation will lead us into the paths of discovery followed by natural scientists of the past is to invoke upon ourselves both obscurity and disappointment. These limitations seem, at present, to be inescapable. And in our dilemma it is perhaps better to secure half a loaf than, through a confusion of our methods, to gain no substantial results at all.

[MS. received September 13, 1929]

DISCUSSION

EXPERIMENTAL EXTINCTION AND NEGATIVE ADAPTATION

In a recent article on inhibition and learning,[1] an attempt was made to show the significance of the extinction process in learning. In the discussion of this form of inhibition types of behavior were discussed which previously had been described as examples of negative adaptation. The author assumed that these forms of behavior represent one general type of activity and that there appeared to be no justification for classifying other types of negative response as different forms of learning. In his revised 'Psychology,'[2] just off the press, Woodworth also discusses such behavior, but classifies it as two different forms of learning. After taking up extinction as a form of learning, he considers negative adaptation under the general topic 'Other Forms of Learning.' The purpose of this paper is to consider Woodworth's argument for such a dichotomy.

The phenomenon of a more or less smoothly progressive weakening of a response to successive unreinforced conditioned stimuli has been called experimental extinction by Pavlov.[3] Woodworth considers this process in his chapter on learning and later as still another form of learning describes negative adaptation as follows: "Apply a harmless and unimportant stimulus time after time in quick succession; at first the animal makes some avoiding or exploring response, but with the continued repetition of the stimulus he ceases to respond, and is said to have become negatively adapted to it. The response has been detached from this particular stimulus."[4] As an example of this form of learning the following experiment with the spider is recounted. "While the spider was on its web, a tuning fork was sounded, and the spider made its regular defensive reaction by dropping to the ground. It climbed back to its web, the fork sounded again, the spider dropped again; but after several repetitions in quick succession, the spider ceased to drop. Next day, to be

[1] A. L. Winsor, Inhibition and learning, PSYCHOL. REV., 1929, 36, 389–401.
[2] R. S. Woodworth, Psychology, rev. ed. 1929.
[3] I. P. Pavlov, Conditioned reflexes, (G. V. Anrep, tr.), Oxford Press, 1927.
[4] R. S. Woodworth, *op. cit.*, p. 161.

sure, it responded as at first; but after several days of this it ceased permanently to respond to the tuning fork."

The similarity of these two forms of learning as described above is most striking. Woodworth recognizes this and adds on the same page that "Negative adaptation is much like the extinction of a conditioned reflex, except that the response that is extinguished in negative adaptation may be a natural and not a conditioned response and except that we have to find the 'check up' that does the extinguishing in negative adaptation." Even if it is true that the natural responses may be extinguished in negative adaptation there would still seem to be no good reason for thinking that a different process is at work in the extinction of an acquired response. The fact that the two types of learning are similar in so many respects would argue for their consideration as the same adaptive process. No such distinction is made between the natural and acquired processes in dealing with the positive aspect of learning.

Confirmation of the identity of these processes may be obtained through an analysis of the physiological theory underlying the extinction process. In the conditioning technique we begin with two organic reactions [5] of different potential strength. The two activities must be repeated simultaneously or the minor stimulus may precede somewhat the major stimulus. The situation as ordinarily illustrated is as follows:

Minor activity—S^1 (metronome) . . . . . R^1 (investigation)
Major activity—S^2 (chewing food) . . . . . R^2 (salivation)

Repetition of this joint activity results in the development of an association between S^1 and R^2, so that excitation of the S^1 elicits a reaction of R^2. Somewhere between the receptor of the minor process and the effector of the major process a dynamic relationship has been established. Presumably this connection has been created somewhere in the cortical or subcortical areas of the brain. The biologically more significant activity has become the policy of the body as a whole. The minor activity assumes the important role of signal agent for the more imperative process. The nervous system has performed its chief function of coordination and integration.

To explain this basic process it is necessary to assume that the functional centers for the concurrent stimuli irradiated nervous energy and that the energy from the weaker center was inducted to-

[5] The common practice of speaking of the 'indifferent' stimulus which becomes conditioned is hardly a true statement of fact. The animal or human is never indifferent to the new stimulus in the beginning of the experiment.

ward the center of stronger potential. R^1 ceases to react to S^1 when the S^2—R^2 process is active and R^2 reacts temporarily to S^1 when S^2 is not present. Whether the blocking of R^1 is an essential feature of the S^1—R^2 relationship is a problem for further research. Pavlov [6] has advanced evidence to show that "induction represents a temporary, phasic phenomenon, which is associated with the period of establishment of new relations in the nervous system; induction makes its appearance only with the maximal development of a given cortical inhibition". . . . It is therefore possible that the flow of energy to R^2 is at least hastened by a blocking between S^1 and R^1. McDougall,[7] on the other hand, suggests that the blocking of R^1 is a complementary result of the increased excitement in the major system, rather than a contributing cause. "Inhibition appears always as the negative or complementary result of a process of increased excitation in some other part." Both theories invite research, but their solution is not essential to our argument. The dynamical processes of irradiation and induction can account for the connection and they may be considered as having been established. Their intimate nature must await further experimentation.

If after this process of conditioning has taken place the metronome is sounded repeatedly without reinforcement (without the accompanying major stimulus) there is more or less progressive extinction of the conditioned response. Nature has again adapted itself to a new situation through the elimination of a useless signal. The determining factor in this adjustment is the prolonged absence of the major stimulus. As long as the major activity occurs contiguously with the minor stimulus the dynamical relationship between S^1 and R^2 can be demonstrated. When the major stimulus ceases to function, extinction of the conditioned response takes place.

Does this description of the extinction process cover also the types of behavior that Woodworth considers as a different form of learning under the term negative adaptation? If it may be assumed that the response of the spider to the tuning fork represents a conditioned reaction the processes are identical. Of the reliability of this assumption there can be little question, as Woodworth himself points out,[8] "Fear, as a response to merely threatened injury, is perhaps always a conditioned response." After many harmless experiences with the tuning fork the fear reaction, the major activity,

[6] I. Pavlov, *op. cit.*, 195.

[7] W. McDougall, The nature of inhibitory processes within the nervous system, *Brain*, 1903, 26, p. 169.

[8] R. S. Woodworth, *op. cit.*, p. 158.

loses its potency and the response to the tuning fork is extinguished just as the salivation to the metronome is extinguished when S^2 disappears. The same analysis might be made of other examples of negative adaptation cited in the chapter under discussion, such as the horse [9] 'getting used' to the harness. The fear reaction evoked at first when the horse was harnessed constituted the major activity. As the fear disappears the extinction of the conditioned response is inevitable, and the harness is no longer a signal for fear. If the extinction process did not take place and the horse continued indefinitely to kick at the harness, we might then be justified in talking about negative adaptation.

It is difficult to understand just what Woodworth means when he insists that we must find the 'check up' that does the extinguishing in negative adaptation.[10] "In the spider's case the extinguishing factor was probably the useless labor of climbing back to its web," he explains. It would seem to us that climbing back to its nest was a very useful activity if the spider was to survive, but the repeated falling process might be represented as useless. According to our description above, however, the 'check up' might well be attributed to the functioning of the more potent innate tendency which was basic in the conditioning process, and this 'check up' is just as important in extinction as in negative adaptation. Psychologists are under the same obligation to 'check up' the causal factors in one type of behavior as another. The need for a 'check up' has doubtful significance as a differentiating characteristic for otherwise similar processes.

As the arguments for differentiating these forms of learning disappear, the necessity for eliminating one or another of the names arises. The key to this process may be found earlier in Woodworth's text. He suggests [11] that "It is a safe rule, then, on encountering any menacing psychological noun, to strip off its linguistic mask and see what manner of activity lies behind." The use of such a term as negative adaptation for the activity described seems particularly 'menacing.' Adaptation means the modifications of an animal or plant fitting it more perfectly for its existence under the conditions of its environment. Negative adaptation then connotes modifications of an animal which fit it *less* perfectly for existence under the conditions of its environment. Certainly 'getting used' to such

[9] Idem, *op. cit.*, p. 162.
[10] Idem, *op. cit.*, p. 162.
[11] Idem, *op. cit.*, p. 6.

factors in the environment as the tuning fork on the part of the spider, or the harness on the part of the horse, could not be considered as literal examples of negative adaptation. Negative adaptation implies abnormal behavior and might well be restricted to such activity. Extinction is not entirely satisfactory, but must remain until a better term is suggested.

The inclusion of the above forms of behavior as types of learning by Woodworth is to be commended. It is a distinct advance from the practice of some psychologists of considering this activity of the spider as a limitation of the law of use. Use, as indicated, is the medium through which the animal adapts himself either in the development of extinction or in the development of the positive reaction. The adjustment of the organism through the elimination of a useless signal is attained by exercise quite as much as the adjustment of the organism through the development of a pertinent signal. These facts deserve consideration in any serious analysis of learning.

A. L. WINSOR

CORNELL UNIVERSITY

[MS. received October 4, 1929]

MEASURES OF MUSICAL TALENT
A REPLY TO DR. C. P. HEINLEIN

Dr. Heinlein has devoted a series of articles to a critical study of my 'Measures of Musical Talent' now available on phonograph records. I desire to comment briefly on some of the findings which he reports, and the principles involved.

In the first place, I wish to express appreciation of the many experimental studies which have been made on these measures. The bibliography is already a long one. In the second place, I wish to say that many of the criticisms brought forth by Dr. Heinlein and other experimenters have been scrutinized and are answered in the rather full critical study of these measures by Dr. Ruth C. Larson, which is now in press and will appear under the title, 'Studies on Seashore's Measures of Musical Talent.' In the third place, Dr. Heinlein's contributions can be evaluated only after contrasting his point of view as one who is against anything that is practical with the effort of one who is trying to render a service. The type of criticism which he makes could be duplicated on any of the best tests that have ever been produced, but without much significance to the testing program.

From the point of view of laboratory research procedure most of his findings are valuable and appropriate, and he shows a real insight into the nature of the musical mind. But he fails to take into account that the measures which he is criticising are practical devices designed for short-cut and snapshot samplings of musical capacities. Now, it is a well-known fact that every test or measure of this kind that has ever been devised, from the Binet test down, represents a compromise, is a practical device and usually represents a fair sampling in a minimum time.

In view of these considerations I shall here merely present my point of view on some issues which he raises. In the study on configuration in tonal memory,[1] he raises three objections to this measure. First, he points out that the elements in each of the five units of the test are not of equal value. My reply is that they were never supposed to be, and it is not necessary that they should be because the test materials are presented in the form of a total block and the scoring is based on the block as a whole. Furthermore, he objects to scoring without weighting for different degrees of difficulty. We conducted experiments for the purpose of making graded units and weighting the score for different degrees of difficulty, but the superiority of that procedure over the method adopted was not sufficient to justify the more elaborate procedure. He objects to a memory span test for musicians. A snapshot test of natural capacities, which is equally fair to musical and unmusical persons for diagnostic purposes is practicable. The tests of achievement, as in memorizing repertoires, would be extremely complicated and would have but little diagnostic value. The most recent experimental studies of this test give it satisfactory reliability and point to a serviceable validity.

In a study on 'The Affective Characters of the Major and Minor Modes in Music'[2] Dr. Heinlein checks the measures on the sense of pitch against the rating of musicians gained by his procedure described in the article and draws the conclusion that pitch discrimination involving as much a difference as five vibrations is not a fundamental requisite in the make-up of the musically trained who have become successful performers. To this I would say that it is all a question of the nature of the performance. The findings might be applicable to the pianist, but if he were dealing with singers, violin-

[1] C. P. Heinlein, A brief discussion of the nature and function of melodic configuration in tonal memory, with critical reference to the Seashore tonal memory test, *J. Genetic Psychol.*, 1928, 35, pp. 45–61.

[2] C. P. Heinlein, *J. Comp. Psychol.*, 1928, 8, pp. 101–142.

ists, and even players of wind instruments, and made a fair correlation, he would find that success in intonation requires at least better than average pitch discrimination, which is three vibrations, and, in the highly skillful, tends to run even below one vibration.

I must here register an objection to the very common practice of validating a measure of this kind against an undifferentiated group. Validation should always be against something which is applicable, and, if possible, specific.

In two studies, 'An Experimental Study of the Seashore Consonance Test,'[3] and 'Critique of the Seashore Consonance Test'[4] Dr. Heinlein brings forth a series of criticisms which are answered in principle by my introductory remarks in this reply. Some of them represent valuable and entirely fitting findings in the study of laboratory procedure. But from the practical point of view I wish to register certain objections. Before doing so, I recognize that the sense of consonance is the least satisfactory of my six measures of musical talent, and many of his objections to the present form would be met by the revision which I propose to make when a new series is imprinted.

The difficulty with the test in its present form is that the instructions are too complicated, requiring the subject to keep three criteria in mind. To eliminate this I propose to simplify the instructions and use only two of the criteria; on one side of the disk, smoothness, and on the other side, blending, selecting intervals in each case so that the criterion shall be unquestionably applicable. Each of these criteria can set a specific task in perception quite isolated from feeling tone in musical progression and would raise the reliability of the test and make it better adapted to children. The instructions would then be, for example, "You will hear two pairs of tones, one combination is smoother than the other. If the second combination is smoother, record S; if rougher, record R." Parallel instructions would be given for blending, if that were chosen for the second criterion.

This change, I am sure, would eliminate several of the criticisms of this test which have been justly made and would make it primarily a test of a particular phase of tonal quality which is involved in the discriminative hearing of consonance and would, in careful procedure, eliminate the element of feeling tone.

Dr. Heinlein points out that in one publication I recognized

[3] C. P. Heinlein. J. Exper. Psychol., 1925, 8, pp. 408–433.

[4] C. P. Heinlein, Psychol. Rev., 1929, 36, pp. 524–542.

purity as one of four criteria, but in the test I used fusion as one criterion. This seems inconsistent, but it is not, because I prefer to use fusion in place of purity on the ground that fusion determines a larger number of units than purity. My objection to fusion as a single basis for determining consonance is the well-recognized fact that the consonance-dissonance array which it reveals does not correspond with any organized order of consonance in certain parts of the array, primarily because rough clangs, as well as smooth ones, tend to fuse. The emphasis on smoothness in the tri-partite instructions, tends to substitute smoothness for fusion in those cases in which roughness is more obvious than non-fusion.

No one denies but that feeling tone is the very essence of consonance-dissonance, as in the preparation for resolution, or even in misanthropic appreciation of dissonance. An extraordinary illustration of the latter influence was found in the case of a musician whom we examined professionally and found that his record on the sense of consonance was 100 per cent negative. A further examination demonstrated that the man was clearly psychopathic and in a negativistic attitude on every issue that confronted him. Can not this extraordinary consistent reversal be counted as equivalent to an equally high record in the right direction? We so interpreted it. I, therefore, agree with Dr. Heinlein that the measurement of consonance in terms of feeling value would be extremely difficult except in so far as the study of one specific factor at a time is undertaken, and can never yield a single order of preference.

I still maintain that the principle of calling for a cognitive judgment is legitimate and will, in careful and unbiased procedures, yield fairly satisfactory results. Experience shows, however, that musically trained people find difficulty in taking the experimental attitude and that for young children the task in its present form is too complicated. For these two reasons very much depends upon the attitude and skill of the experimenter, as is well illustrated in measures of reliability which have been published.

It may be said, then, that the more experienced the experimenter is in the handling of group testing, and the freer he is from prejudice and bias, the higher the reliability of the test will be found. For adults, Dr. Ruth Larson got r. 71, Miss Delia Larson r .63, Professor Peterson r .68, Dr. Gaw r .57, and Salisbury and Smith got r .69, all with satisfactory probable error; while for Brown, who got r .43 and Weaver, who got r .35, there were undoubtedly specific sources of error in the procedure, as may also be true in the case of Heinlein,

for their lower correlations. While correlations of between 60 and 70 are low, they are significant. But it is unfortunate in this test, as in most others, that the reliability is to some extent dependent upon the experimenter. For children, Dr. Ruth Larson found r .54 for the eighth grade, r .57 for the seventh grade, r .57 for the sixth grade, and r .36 for the fifth grade. These correlations are not satisfactory and for that reason this test in the present form should be used with reserve on children.

Those who criticize the validity of the test often forget that the test is specific and that the validity should be tested against performance in which this factor functions. The mistake is often made of assuming that persons who have had music lessons, or even are engaged professionally in music, therefore have a good sense of consonance. For this reason correlations of the type of Heinlein's and Highsmith's are of little significance. On the other hand, even with its present weakness or low reliability, this measure, when used in a battery to produce a general profile, does contribute to some extent toward an index of capacity for success in higher music.

Dr. Heinlein calls for proof to show that this measure is elemental. There are now a number of studies which show that when suitable allowance is made for the conditions of testing, this particular measure varies very little with age and with training; and that is what I mean when I speak of this and other measures as being relatively elemental. This, of course, does not necessarily involve any assumption in regard to the mechanism of specificity in inheritance and is always qualified by the difference in test response at different ages. The very slight intercorrelation between this measure and the other five measures in this series is good support for including this particular unit in a battery since it measures something different from what is contained in each of the other measures.

It is claimed that I regard consonance as static rather than dynamic. I regard consonance, as such, in the musical situation as dynamic. But in the test procedure here discussed, I maintain that two clangs may be taken out for comparison in a static situation, provided the instructions call for a cognitive judgment, and in such a situation can be made clear and effective.

The use of the phonograph is criticized, but we must bear in mind that for the purpose of standardizing, the norms are based upon records from the phonograph, and by this instrument we probably obtain more constant conditions than we could by any other instrument available under the test conditions. The claim that the

slight preliminary practice, recommended for the purpose of explaining the procedures, becomes vitiating training for the purpose of the test, I do not take seriously, because it is very small in comparison with the whole series, and no one has any way of knowing which stimuli that follow later were used in the demonstration on any other instrument. It is true that variations in intensity affect judgments of this kind. In recording, the players of the horns were asked to make all tones equally loud, and that distortion which comes from the recording instruments is fairly taken into account by our standardizing of the test on the phonograph records instead of on the original tones.

In general, then, it is admitted that this measure is less satisfactory than any of the other five in that battery, and that a decided improvement can be made along the line suggested above. But at the best, any test is full of sources of error due to a necessary compromise between the theoretical and the practical, and no amount of refinement of a test of this sort is likely to satisfy one who seems to assume that no practical test is possible.

CARL E. SEASHORE

STATE UNIVERSITY OF IOWA

[MS. received December 4, 1929]

A REPLY TO PROFESSOR BENTLEY

Professor Bentley's "Note on 'Observer' and 'Subject'" appearing in the October issue of the *American Journal of Psychology*, is quite obviously directed at my "Note on Use of the Term 'Observer'" appearing in the November issue of the PSYCHOLOGICAL REVIEW, which he had seen in manuscript some months ago,—remarking at the time that my note should have editorial comment, inasmuch as my proposal "overlooked quite definite distinctions made for some time between the use of the terms 'subject' and 'observer.'" In further correspondence I failed to learn what distinctions were meant; but Professor Bentley has been explicit in his published note which successfully anticipated mine.

Since Professor Bentley's remarks may lead the reader to a misinterpretation of my own Note, I beg leave to call attention to two or three points. For one thing, he evidently misunderstands my intent when he speaks of "the persistent error of the 'objectivist' who naïvely insists upon identifying all reports of observers with 'subjectivism' and 'introspectionism,'" when he is at pains to show

the shortcomings of a "cult of objectivism," and so on for the major portion of his article. Now it happens that a part of the argument in my Note is couched in these phrases: "Originally the word ['observer'] was in use to designate the person whose experiences were the subject matter of observation. . . . But by no means are all the problems of psychology concerned with a person's (the 'subject's') direct experience. . . . " If I must point it out explicitly—it had seemed to me that the plain implication of such words was: neither are all the problems of psychology limited to the subject's behavior as observed and recorded by another person. If I have insight into my own motivation, my Note was not written in furtherance of objectivism or any other 'ism nor in criticism of subjectivism or any other 'ism.

An error of fact appears to underlie an assumption tacitly made in Professor Bentley's Note; namely, that to deprive the psychologist of the use of the term of 'observer' in the sense I criticized would be to cripple him seriously. To my Note, after Professor Bentley had seen it in manuscript, I added a sentence listing eleven experimental manuals in which the term 'subject' is used where some others (as Titchener) would have employed 'observer'—and no one would accuse the compilers of those manuals (Sanford, C. S. Myers, *et al.*) of any objectivistic bias! The error would seem to be one of assuming that what happens to be a particular terminology employed by Titchener and his disciples is the only one actually in use by other psychologists working upon precisely the same problems.

Finally, there is an incidental inference one may draw from Professor Bentley's Note which calls for examination because of its rather general acceptance. Whether or not he means to imply just this, there is abroad the opinion that the psychological 'observer' (in Titchener's sense) and his psychological 'observations' are similar in principle and play the same scientific rôle as do the observer and his observations in any other natural science. This I question. Observing by an observer in astronomy, physics, or geology, is designed to get at the structures and functioning of entities and agencies conceived as essentially independent of the observer—although gotten at through him; whereas in psychology the emphasis is shifted to entities and agencies *as* observed and while *under* observation. Even when the peculiarities of the particular individuals doing the observing are cancelled out in order to arrive at norms or laws or general principles (such as, *e.g.*, the three laws of color mix-

ture), these latter are yet normal or generalized observations *qua* observations. In other words, the emphasis always remains upon the human observer; and the objective of the study is (as Titchener puts it neatly) "to describe the world . . . as it appears with man left in." I hold, then, that it is inaccurate to insist that the 'observer' in psychology plays a rôle as similar to that played by the astronomical observer, say, as the rôle of the latter is to that of the geological or entomological. Even when the *Versuchsperson* is furnishing a record of his experiences in cutaneous sensibility, in visual after imagery, etc., these are and will always be considered as personal rather than impersonal data.

In this connection, it is to be noted that when the *Versuchsperson* is to make his own observations on himself, is to describe his own experiences (or word it as you will), these *his* experiences form the *subject*-matter of the experiment. It follows, therefore, that there should be no impropriety in referring to the person concerned as the 'subject.' Whether the experiment be conducted entirely by another person or entirely by himself or by both, he is there personally to furnish the subject-matter in his own personal way; *i.e.*, he is the subject of the experiment!

And now, let me re-state my original recommendation. Whereas confusion often arises from the use of the two different words, 'observer' and 'subject,' to refer to the person who (or whose experience, or both) is under observation (especially in those many studies where the same person is studied both objectively and subjectively), an interest in clarity would demand the use of one common word in all cases; and, whereas 'observer' simply cannot be used in all cases (for the *Versuchsperson* does not always observe) and whereas 'subject' can be so used without impropriety (as was above indicated), I recommend the universal use of the term 'subject,' as one that would be accurate enough whether the *Versuchsperson* is being studied objectively or subjectively.

UNIVERSITY OF NORTH CAROLINA

J. F. DASHIELL

[MS. received December 9, 1929]

Psychological Review Publications

Original contributions and discussions intended for the Psychological Review should be addressed to

Professor Howard C. Warren, Editor Psychological Review,
Princeton University, Princeton, N. J.

Original contributions and discussions intended for the Journal of Experimental Psychology should be addressed to

Professor Madison Bentley, Editor Journal of Experimental Psychology,
Morrill Hall, Cornell University, Ithaca, N. Y.

Contributions intended for the Psychological Monographs should be addressed to

Professor Raymond Dodge, Editor Psychological Monographs,
Kent Hall, Yale University, New Haven, Conn.

Reviews of books and articles intended for the Psychological Bulletin, announcements and notes of current interest, and *books offered for review* should be sent to

Professor S. W. Fernberger, Editor Psychological Bulletin,
University of Pennsylvania, Philadelphia, Pa.

Titles and reprints intended for the Psychological Index should be sent to

Professor Walter S. Hunter, Editor Psychological Index,
Clark University, Worcester, Mass.

All business communications should be addressed to
Psychological Review Company, Princeton, N. J.

The Psychological Review
is indexed in the
International Index to Periodicals
to be found in most public and
college libraries

DIRECTORY OF
American Psychological Periodicals

American Journal of Psychology—Ithaca, N. Y.; Cornell University. Subscription $6.50. 624 pages ann. Ed. by M. F. Washburn, K. M. Dallenbach, Madison Bentley and E. G. Boring. Quarterly. General and experimental psychology. Founded 1887.

The Pedagogical Seminary and Journal of Genetic Psychology—Worcester, Mass.; Clark University Press. Subscription $7.00. 700 pages ann. Ed. by Carl Murchison. Quarterly. Child behavior, differential and genetic psychology. Founded 1891.

Psychological Review—Princeton, N. J.; Psychological Review Company. Subscription $5.50. 500 pages annually. Bi-monthly. General. Founded 1894. Edited by Howard C. Warren.

Psychological Monographs—Princeton, N. J.; Psychological Review Company. Subscription $6.00 per vol. 500 pp. Founded 1895. Ed. by Shepherd I. Franz. Published without fixed dates, each issue one or more researches.

Psychological Index—Princeton, N. J.; Psychological Review Company. Subscription $4.00. 300–400 pp. Founded 1895. Edited by W. S. Hunter. An annual bibliography of psychological literature.

Journal of Philosophy—New York; 515 W. 116th Street. Subscription $4. 728 pages per volume. Founded 1904. Bi-weekly. Edited by F. J. E. Woodbridge, Wendell T. Bush and H. W. Schneider.

Psychological Bulletin—Princeton, N. J.; Psychological Review Company. Subscription $6.00. 720 pages annually. Psychological literature. Monthly. Founded 1904. Edited by Samuel W. Fernberger.

Training School Bulletin—Vineland, N. J.; The Training School. Subscription $1. 160 pages ann. Ed. by E. R. Johnstone. Founded 1904. Monthly (10 numbers). Psychology and training of defectives.

Archives of Psychology—Columbia University P. O., New York City. Subscription $6. 500 pp. per vol. Founded 1906. Ed. by R. S. Woodworth. Published without fixed dates, each number a single experimental study.

Journal of Abnormal Psychology and Social Psychology—Eno Hall, Princeton, N. J. Subscription $5. Acting Editor, Henry T. Moore. Quarterly. 432 pages ann. Founded 1906. Abnormal and social.

Psychological Clinic—Philadelphia; Psychological Clinic Press. Subscription $3.00. 288 pages. Ed. by Lightner Witmer. Founded 1907. Without fixed dates (9 numbers). Orthogenics, psychology, hygiene.

Psychoanalytic Review—Washington, D. C.; 3617 10th St., N. W. Subscription $6. 500 pages annually. Psychoanalysis. Quarterly. Founded 1913. Ed. by W. A. White and S. E. Jelliffe.

Journal of Experimental Psychology—Princeton, N. J. Psychological Review Company. 500 pages annually. Experimental. Subscription $6.00. Founded 1916. Bi-monthly. Ed. by Madison Bentley.

Journal of Applied Psychology—Baltimore, Md.; Williams & Wilkins Co. Subscription $4. 400 pages annually. Founded 1917. Quarterly. Edited by James P. Porter and William F. Book.

Journal of Comparative Psychology—Baltimore; Williams & Wilkins Company. Subscription $5. 500 pages annually. Founded 1921. Bi-monthly. Edited by Knight Dunlap and Robert M. Yerkes.

Comparative Psychology Monographs—Baltimore; Williams & Wilkins Co. Subscription $5. 500 pages per volume. Edited by W. S. Hunter. Published without fixed dates, each number a single research. Founded 1922.

Genetic Psychology Monographs—Worcester, Mass.; Clark University Press. Subscription $7.00 per volume. Two volumes per year, 600 pages each. Ed. by Carl Murchison. Monthly. Each number one complete research. Child behavior, differential and genetic psychology. Founded 1925.

Psychological Abstracts—Eno Hall, Princeton, N. J. Edited by W. S. Hunter. Subscription $6.00. Monthly. 700 pages annually. Founded 1927.

Journal of General Psychology—Worcester, Mass. Clark University Press. Subscription $7.00. 600–700 pages annually. Edited by Carl Murchison. Quarterly. Experimental, theoretical, clinical, and historical psychology. Founded 1927.

Journal of Social Psychology—Worcester, Mass. Clark University Press. Subs. $7.00. 500–600 pp. annually. Edited by John Dewey, Carl Murchison and internat'l cooperating board. Founded 1929. Quarterly. Political, racial, and differential psychology.

PH.